(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 December 2003 (31.12.2003)

PCT

(10) International Publication Number
WO 2004/001058 A2

(51) International Patent Classification[7]: C12Q

(21) International Application Number: PCT/US2002/014255

(22) International Filing Date: 6 May 2002 (06.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/288,660 4 May 2001 (04.05.2001) US

(71) Applicant *(for all designated States except US)*: **PARATEK PHARMACEUTICALS, INC.** [US/US]; 75 Kneeland Street, Boston, MA 02111 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEVY, Stuart, B.** [US/US]; 144 Warren Avenue, Boston, MA 02116 (US). **ALEKSHUN, Micheal, N.** [US/US]; 44 Hawley Road, Wakefield, MA 01880 (US). **PODLOGAR, Brent** [US/US]; 18 Quarry Lane, Hamden, CT 06518 (US). **OHEMENG, Kwasi** [US/US]; 400 Engamore Lane, Apt 101, Norwood, MA 02062 (US). **VERMA, Atul** [IN/US]; 1294 Massachusetts Avenue, Arlington, MA 02476 (US). **WARCHOL, Tadeusz** [PL/US]; 411 Great Road, Acton, MA 01720 (US). **BHATIA, Beena** [IN/US]; 1294 Massachusetts Avenue, Arlington, MA 02476 (US).

(74) Agents: **HANLEY, Elizabeth, A.** et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** TRANSCRIPTION FACTOR MODULATING COMPOUNDS AND METHODS OF USE THEREOF

(57) **Abstract:** Methods for identifying compound useful as anti-infectives that decrease resistance, virulence, or growth of microbes are provided. In one embodiment, the method comprises contacting a microbial cell comprising: 1) a selectable marker under the control of a transcription factor responsive element and 2) a transcription factor, with a compound under conditions which allow interaction of the compound with the microbial cell; and measuring the ability of the compound to affect the growth or survival of the microbial cell as an indication of whether the test compound modulates the activity of a transcription factor.

WO 2004/001058 A2

TRANSCRIPTION FACTOR MODULATING COMPOUNDS AND METHODS OF USE THEREOF

Related Applications

This application claims priority to U.S. Provisional Application Serial No. 60/288,660, entitled "Helix-Turn-Helix Protein Modulating Compounds and Methods of Use Thereof," filed on May 4, 2001, the entire contents of which are hereby incorporated herein by reference.

Background of the Invention

Multidrug resistance in bacteria is generally attributed to the acquisition of multiple transposons and plasmids bearing genetic determinants for different mechanisms of resistance (Gold *et al.* 1996. *N. Engl. J. Med.* 335:1445). However, descriptions of intrinsic mechanisms that confer multidrug resistance have begun to emerge. The first of these was a chromosomally encoded multiple antibiotic resistance (*mar*) locus in *Escherichia coli* (George and Levy, 1983. *J. Bacteriol.* 155:531; George and Levy 1983 *J. Bacteriol.* 155:541). Mar mutants of *E. coli* arose at a frequency of 10^{-6} to 10^{-7} and were selected by growth on subinhibitory levels of tetracycline or chloramphenicol (George and Levy, supra). These mutants exhibited resistance to tetracyclines, chloramphenicol, penicillins, cephalosporins, puromycin, nalidixic acid, and rifampin (George and Levy, supra). Later, the resistance phenotype was extended to include fluoroquinolones (Cohen et al. 1989. *Antimicrob. Agents Chemother.* 33:1318), oxidative stress agents (Ariza et al. 1994. *J. Bacteriol.* 176:143; Greenberg *et al.* 1991. *J. Bacteriol.* 73:4433), and more recently, organic solvents (White *et al.* 1997. *J. of Bacteriology* 179:6122; Asako, *et al.* 1997. *J. Bacteriol.* 176:143) and household disinfectants, *e.g.*, pine oil and/or TRICLOSAN® (McMurry *et al.* 1998. *FEMS Microbiology Letters* 166:305; Moken *et al.* 1997. *Antimicrobial Agents and Chemotherapy* 41:2770).

The *mar* locus consists of two divergently positioned transcriptional units that flank a common promoter/operator region in *E. coli*, *Salmonella typhimurium*, and other *Entrobacteriacae* (Alekshun and Levy. 1997, *Antimicrobial Agents and Chemother.* 41: 2067). One operon encodes MarC, a putative integral inner membrane protein without any yet apparent function, but which appears to contribute to the Mar phenotype in some strains. The other operon comprises *marRAB*, encoding the Mar repressor (MarR), which binds *marO* and negatively regulates expression of *marRAB* (Cohen et al. 1994. *J. Bacteriol.* 175:1484; Martin and Rosner 1995. *PNAS* 92:5456; Seoane and Levy. 1995 *J. Bacteriol.* 177:530), an activator (MarA), which controls expression of other genes on the chromosome, *e.g.*, the *mar* regulon (Cohen et al. 1994

J. Bacteriol. 175:1484; Gambino et. al. 1993. *J. Bacteriol.* 175:2888; Seoane and Levy, 1995 *J. Bacteriol.* 177:530), and a putative small protein (MarB) of unknown function.

Exposure of *E. coli* to several chemicals, including tetracycline and chloramphenicol (Hachler *et al.* 1991 *J Bacteriol* 173(17):5532-8; Ariza, 1994, *J Bacteriol*; 176(1):143-8), sodium salicylate and its derivatives (Cohen, 1993, *J Bacteriol*; 175(24):7856-62) and oxidative stress agents (Seoane *et al.* 1995. *J Bacteriol*; 177(12):3414-9) induces the Mar phenotype. Some of these chemicals act directly at the level of MarR by interacting with the repressor and inactivating its function (Alekshun. 1999. *J. Bacteriol.* 181:3303-3306) while others (antibiotics such as tetracycline and chloramphenicol) appear to induce mar expression by an alternative mechanism (Alekshun. 1999. *J. Bacteriol.* 181:3303-3306) *e.g.*, through a signal transduction pathway.

Once expressed, MarA activates the transcription of several genes that constitute the E. coli mar regulon (Alekshun, 1997, *Antimicrob. Agents Chemother.* 41:2067-2075; Alekshun, 1999, *J. Bacteriol.* 181:3303-3306). With respect to decreased antibiotic susceptibility, the increased expression of the AcrAB/TolC multidrug efflux system (Fralick, 1996, *J Bacteriol.* 178(19):5803-5; Okusu, 1996 J Bacteriol;178(1):306-8) and decreased synthesis of OmpF (Cohen, 1988, *J Bacteriol.*; 170(12):5416-22) an outer membrane protein, play major roles. Organic solvent tolerance, however, is attributed to MarA mediating increased expression of AcrAB, TolC, OmpX, and a 77 kDa protein (Aono, 1998, *Extremophiles*; 2(3):239-48; Aono, 1998 *J Bacteriol*; 180(4):938-44.) but is independent of OmpF levels (Asako, 1999, *Appl Environ Microbiol*; 65(1):294-6).

MarA is a member of the XylS/AraC family of transcriptional activators (Gallegos et al. 1993. *Nucleic Acids Res.* 21:807). There are more than 100 proteins within the XylS/AraC family and a defining characteristic of this group of proteins is the presence of two helix-turn-helix (HTH) DNA binding motifs. Proteins within this family activate many different genes, some of which produce antibiotic and oxidative stress resistance or control microbial metabolism and virulence (Gallegos et al. supra).

Summary of the Invention

The present invention represents an advance over the prior art by identifying transcription factor modulating compounds, such as, but not limited to helix-turn-helix protein modulating compounds, and providing novel assays that can be used to identify compounds which modulate microbial transcription factors, such as MarA family polypeptides and AraC family polypeptides. Modulation of gene transcription brought about by the modulation of transcription factors, such as helix-turn-helix domain containing proteins, can control a wide variety of cellular processes. For

example, in prokaryotic cells processes such as metabolism, resistance, and virulence can be controlled.

Assays to identify compounds that are capable of modulating bacterial transcription factors would be of great benefit in the identification of agonists and antagonists that can be used to control gene transcription in both prokaryotic and eukaryotic cells.

In one embodiment, the invention pertains to a method for reducing antibiotic resistance of a cell, e.g., a eukaryotic or prokaryotic cell. In a preferred embodiment, the cell is a microbial cell. In one embodiment, the invention pertains to a method for reducing antibiotic resistance in a microbial cell, by contacting a cell with a transcription factor modulating compound, such that the antibiotic resistance of the cell is reduced. In an embodiment, the transcription factor modulating compound is of the formula (I):

A-E (I)

wherein A is a polar moiety; E is a hydrophobic moiety, and pharmaceutically acceptable salts thereof.

In another embodiment, the invention pertains to a method for modulating transcription. The method includes contacting a transcription factor with a transcription factor modulating compound, such that the transcription factor is modulated. The transcription factor modulating compound is of the formula (I):

A-E (I)

wherein A is a polar moiety; and E is a hydrophobic moiety, and pharmaceutically acceptable salts thereof.

In another embodiment, the invention also includes methods for identifying transcription factor modulating compounds. The method includes contacting a microbial cell with a test compound under conditions which allow interaction of the compound with the microbial cell and measuring the ability of the test compound to affect the cell. The microbial cell includes a selective marker under the direct control of a transcription factor responsive element and a transcription factor.

In yet another embodiment, the invention includes methods for identifying a transcription factor modulating compound. The method includes contacting a microbial cell comprising: 1) a selective marker under the control of a transcription factor responsive element and 2) a transcription factor, with a test compound under conditions which allow interaction of the compound with the microbial cell, and measuring the ability of the test compound to affect the growth (*e.g.*, *in vitro* or *in vivo*) or survival of the microbial cell, wherein the inactivation of the transcription

factor leads to a decrease in *in vitro* or *in vivo* cell survival. The invention also pertains to similar methods where the inactivation of the transcription factor leads to an increase in cell survival, as well as methods wherein the activation of the transcription factor leads to increased or, alternatively, decreased cell survival.

In another embodiment, the invention also pertains to methods for identifying a transcription factor modulating compound, by contacting a microbial cell comprising: 1) a chromosomal deletion in a guaB or purA gene, 2) heterologous guaB or purA gene under the control of its transcription factor responsive promoter, and 3) a transcription factor, with a test compound under conditions which allow interaction of the compound with the microbial cell. The method further includes the steps of measuring the ability of the compound to affect gene expression of the reporter or the growth or survival of the microbial cell as an indication of whether the compound modulates the activity of a transcription factor. The ability of the compound to modulate the activity of a transcription factor leads to an alteration in gene expression may effect cell growth or survival.

The invention pertains to transcription factor modulating compounds, HTH protein modulating compounds, and MarA family modulating compounds identified by the methods of the invention, methods of using these compounds and pharmaceutical compositions comprising these compounds. The transcription factor modulating compounds of the invention include, but are not limited to, compounds of formulae (I)-(X) and Tables 4 and 5.

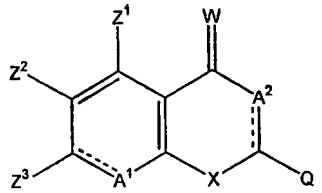
The invention also pertains to methods using computer modeling programs to identify transcription factor modulating compounds. For example, the invention pertains to a method of identifying transcription factor modulating compounds. The method includes obtaining the structure of the transcription factor modulating compound, and using or identifying a scaffold which has an interaction energy score of –20 or less with a portion of the transcription factor, thus identifying potential transcription factor modulating scaffolds.

The invention also pertains, at least in part, to a kit for identifying a transcription factor modulating compound which modulates the activity of a transcription factor polypeptide comprising a microbial cell. The kit includes 1) a selective marker under the control of a transcription factor responsive element and 2) a transcription factor.

The invention also pertains, at least in part, to pharmaceutical compositions which contain an effective amount of a transcription factor modulating compound, and, optionally, a pharmaceutically acceptable carrier.

The invention also pertains to a method of inhibiting a biofilm, by administering a composition comprising a transcription factor modulating compound such that the biofilm is inhibited.

In a further embodiment, the invention pertains to a pharmaceutical composition comprising an effective amount of a transcription factor modulating compound, and a pharmaceutically acceptable carrier. The transcription factor modulating compound is of the formula (II):



(II)

wherein

W is O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^4, O, or S;

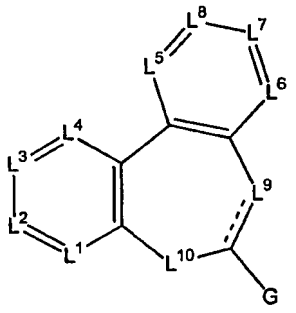
A^2 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, Z^4 and Z^5 are each independently hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, or cyano;

Z^3 is hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, nitro, cyano, carbonyl, or thiocarbonyl;

Q is an aromatic or heterocyclic moiety, and pharmaceutically acceptable salts thereof.

In another further embodiment, the invention pertains to a pharmaceutical composition comprising an effective amount of a transcription factor modulating compound, and a pharmaceutically acceptable carrier. The compound is of the formula (III):



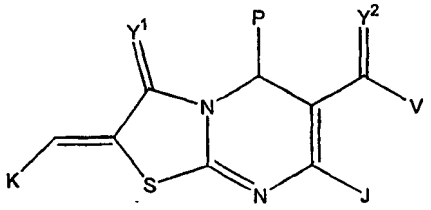
(III)

wherein

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen; and

L^1, L^2, L^3, L^4, L^5, L^6, L^7, L^8, L^9, and L^{10} are each independently oxygen, substituted or unsubstituted nitrogen, sulfur and or substituted or unsubstituted carbon, and pharmaceutically acceptable salts thereof.

In yet another embodiment, the invention pertains to a pharmaceutical composition comprising an effective amount of a transcription factor modulating compound and a pharmaceutically acceptable carrier. The transcription factor modulating compound is of the formula (IV):



(IV)

wherein

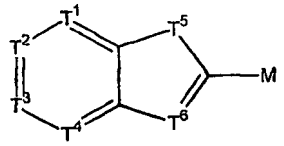
Y^1 and Y^2 are each oxygen or sulfur;

J is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, cyano, nitro, amino, or halogen;

V is substituted or unsubstituted alkyl, alkenyl, alkynyl, alkoxy, alkylamino, or alkylthio;

P and K are each independently substituted or unsubstituted aryl, and pharmaceutically acceptable salts thereof.

In another embodiment, the invention pertains to a pharmaceutical composition comprising a pharmaceutically acceptable carrier and a transcription factor modulating compound. The transcription factor modulating compound is of the formula (V):



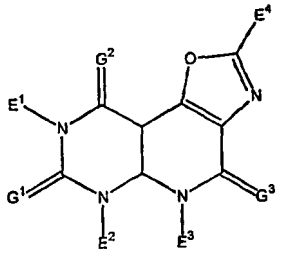
(V)

wherein

T^1, T^2, T^3, T^4, T^5, and T^6 are each independently substituted or unsubstituted carbon, oxygen, substituted or unsubstituted nitrogen, or sulfur;

M is hydrogen, alkyl, alkenyl, heterocyclic, alkynyl, or aryl, or pharmaceutically acceptable salts thereof.

In another embodiment, the invention pertains to a pharmaceutical composition comprising a pharmaceutically acceptable carrier and a transcription factor modulating compound. The transcription factor modulating compound may be of the formula (VI):



(VI)

wherein

G^1, G^2, and G^3 are each independently O, S, substituted or unsubstituted nitrogen, or substituted or unsubstituted carbon;

E^1, E^2, and E^3 are each independently hydrogen, alkyl, alkenyl, alkynyl, aryl, aralkyl, or acyl; and

E^4 is alkyl, alkenyl, alkynyl, aryl, halogen, cyano, amino, nitro, or acyl, and pharmaceutically acceptable salts thereof.

Brief Description of the Drawings

Figure 1 is a multiple sequence alignment of AraC-XylS family polypeptides.

Figure 2 is a multiple sequence alignment of PROSITE PS00041 and AraC family polypeptides.

Figure 3 is a multiple sequence alignment of PROSITE PS01124 and AraC family polypeptides.

Figure 4 is a CoMFA contour map for a representative triazinoxazepine.

Detailed Description of the Invention

The invention pertains, at least in part, to compounds which modulate transcription factors (e.g., helix-turn-helix (HTH) proteins, AraC family polypeptides, MarA family polypeptides, etc.), methods of identifying the transcription factor modulating compounds (e.g., HTH protein modulating compounds, AraC family polypeptide modulating compounds, MarA family polypeptide modulating compounds, etc.), and methods of using the compounds.

1. Transcription factors

The term "transcription factor" includes proteins that are involved in gene regulation in both prokaryotic and eukaryotic organisms. In one embodiment, transcription factors can have a positive effect on gene expression and, thus, may be referred to as an "activator" or a "transcriptional activation factor." In another embodiment, a transcription factor can negatively effect gene expression and, thus, may be referred to as "repressors" or a "transcription repression factor." Activators and

repressors are generally used terms and their functions are discerned by those skilled in the art.

The term "AraC family polypeptide," "AraC-XylS family polypeptide" or "AraC-XylS family peptide" include an art recognized group of prokaryotic transcription factors which contains more than 100 different proteins (Gallegos *et al.*, (1997) *Micro. Mol. Biol. Rev.* 61: 393; Martin and Rosner, (2001) *Curr. Opin. Microbiol.* 4:132). AraC family polypeptides include proteins defined in the PROSITE (PS) database (http://www.expasy.ch/prosite/) as profile PS01124. The AraC family polypeptides also include polypeptides described in PS0041, HTH AraC Family 1, and PS01124, and HTH AraC Family 2. Multiple sequence alignments for the AraC-XylS family polypeptides, HTH AraC family 1, and HTH AraC family 2 are shown in Figures 1-3, respectively. In an embodiment, the AraC family polypeptides are generally comprised of, at the level of primary sequence, by a conserved stretch of about 100 amino acids, which are believed to be responsible for the DNA binding activity of this protein (Gallegos *et al.*, (1997) *Micro. Mol. Biol. Rev.* 61: 393; Martin and Rosner, (2001) *Curr. Opin. Microbiol.* 4: 132). AraC family polypeptides also may include two helix turn helix DNA binding motifs (Martin and Rosner, (2001) *Curr. Opin. Microbiol.* 4: 132; Gallegos *et al.*, (1997) *Micro. Mol. Biol. Rev.* 61: 393; Kwon *et al.*, (2000) *Nat. Struct. Biol.* 7: 424; Rhee *et al.*, (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95: 10413). The term includes MarA family polypeptides and HTH proteins. In one embodiment, the invention pertains to a method for modulating an AraC family polypeptide, by contacting the AraC family polypeptide with a test compound which interacts with a portion of the polypeptide involved in DNA binding. In a further embodiment, the test compound interacts with a conserved aminoacid residue (capitalized) of the HTH AraC family 1 protein indicated in Figure 2.

The term "helix-turn-helix protein," "HTH protein," "helix-turn-helix polypeptides," and "HTH polypeptides," includes proteins comprising one or more helix-turn-helix domains. Helix-turn-helix domains are known in the art and have been implicated in DNA binding (*Ann Rev. of Biochem.* 1984. 53:293). An example of the consensus sequence for a helix-turn domain can be found in Brunelle and Schleif (1989. *J. Mol. Biol.* 209:607). The domain has been illustrated by the sequence XXXPhoAlaXXPhoGlyPhoXXXXPhoXXPhoXX, where X is any amino acid and Pho is a hydrophobic amino acid.

The helix-turn-helix domain was the first DNA-binding protein motif to be recognized. Although originally the HTH domain was identified in bacterial proteins, the HTH domain has since been found in hundreds of DNA-binding proteins from both

eukaryotes and prokaryotes. It is constructed from two alpha helices connected by a short extended chain of amino acids, which constitutes the "turn."

In one embodiment, a helix-turn-helix domain containing protein is a Mar A family polypeptide. The language "MarA family polypeptide" includes the many naturally occurring HTH proteins, such as transcription regulation proteins which have sequence similarities to MarA and which contain the MarA family signature pattern, which can also be referred to as an XylS/AraC signature pattern. An exemplary signature pattern which defines MarA family polypeptides is shown, *e.g.*, on PROSITE and is represented by the sequence: [KRQ]-[LIVMA]-X(2)-[GSTALIV]-{FYWPGDN}X(2)-[LIVMSA]-X(4,9)-[LIVMF]-X(2)-[LIVMSTA]-X(2)-[GSTACIL]-X(3)-[GANQRF]-[LIVMFY]-X(4,5)-[LFY]-X(3)-[FYIVA]-{FYWHCM}-X(3)-[GSADENQKR]-X-[NSTAPKL]-[PARL], where X is any amino acid. MarA family polypeptides have two "helix-turn-helix" domains. This signature pattern was derived from the region that follows the first, most amino terminal, helix-turn-helix domain (HTH1) and includes the totality of the second, most carboxy terminal helix-turn-helix domain (HTH2). (See PROSITE PS00041).

The MarA family of proteins ("MarA family polypeptides") represent one subset of AraC-XylS family polypeptides and include proteins like MarA, SoxS, Rob, Rma, AarP, PqrA, etc. The MarA family polypeptides, generally, are involved in regulating resistance to antibiotics, organic solvents, and oxidative stress agents (Alekshun and Levy, (1997) *Antimicrob. Agents. Chemother.* 41: 2067). Like other AraC-XylS family polypeptides, MarA-like proteins also generally contain two HTH motifs as exemplified by the MarA and Rob crystal structures (Kwon *et al.*, (2000) *Nat. Struct. Biol.* 7: 424; Rhee *et al.*, (1998) *Proc. Natl. Acad. Sci.* U.S.A. 95: 10413). Members of the MarA family can be identified by those skilled in the art and will generally be represented by proteins with homology to amino acids 30-76 and 77-106 of MarA (SEQ ID. NO. 1).

Preferably, a MarA family polypeptide or portion thereof comprises the first MarA family HTH domain (HTH1) (Brunelle, 1989, *J Mol Biol*; 209(4):607-22). In another embodiment, a MarA polypeptide comprises the second MarA family HTH domain (HTH2) (Caswell, 1992, *Biochem J.*; 287:493-509.). In a preferred embodiment, a MarA polypeptide comprises both the first and second MarA family HTH domains.

MarA family polypeptide sequences are "structurally related" to one or more known MarA family members, preferably to MarA. This relatedness can be shown by sequence or structural similarity between two MarA family polypeptide sequences or between two MarA family nucleotide sequences that specify such polypeptides. Sequence similarity can be shown, *e.g.*, by optimally aligning MarA family member sequences using an alignment program for purposes of comparison and comparing

corresponding positions. To determine the degree of similarity between sequences, they will be aligned for optimal comparison purposes (*e.g.*, gaps may be introduced in the sequence of one protein for nucleic acid molecule for optimal alignment with the other protein or nucleic acid molecules). The amino acid residues or bases and corresponding amino acid positions or bases are then compared. When a position in one sequence is occupied by the same amino acid residue or by the same base as the corresponding position in the other sequence, then the molecules are identical at that position. If amino acid residues are not identical, they may be similar. As used herein, an amino acid residue is "similar" to another amino acid residue if the two amino acid residues are members of the same family of residues having similar side chains. Families of amino acid residues having similar side chains have been defined in the art (see, for example, Altschul *et al.* 1990. *J. Mol. Biol.* 215:403) including basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan). The degree (percentage) of similarity between sequences, therefore, is a function of the number of identical or similar positions shared by two sequences (*i.e.*, % homology = # of identical or similar positions/total # of positions x 100). Alignment strategies are well known in the art; see, for example, Altschul et al. *supra* for optimal sequence alignment.

MarA family polypeptides may share some amino acid sequence similarity with MarA. The nucleic acid and amino acid sequences of MarA as well as other MarA family polypeptides are available in the art. For example, the nucleic acid and amino acid sequence of MarA can be found, *e.g.*, on GeneBank (accession number M96235 or in Cohen *et al.* 1993. *J. Bacteriol.* 175:1484, or in SEQ ID NO:1 and SEQ ID NO:2.

The nucleic acid and/or amino acid sequences of MarA can be used as "query sequences" to perform a search against databases (*e.g.*, either public or private) to, for example, identify other MarA family members having related sequences. Such searches can be performed, *e.g.*, using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to MarA family nucleic acid molecules. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to MarA protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can

be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

MarA family members can also be identified as being similar based on their ability to specifically hybridize to nucleic acid sequences specifying MarA. Such stringent conditions are known to those skilled in the art and can be found *e.g.*, in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Conditions for hybridizations are largely dependent on the melting temperature Tm that is observed for half of the molecules of a substantially pure population of a double-stranded nucleic acid. Tm is the temperature in °C at which half the molecules of a given sequence are melted or single-stranded. For nucleic acids of sequence 11 to 23 bases, the Tm can be estimated in degrees C as 2(number of A+T residues) + 4(number of C+G residues). Hybridization or annealing of nucleic acid molecules should be conducted at a temperature lower than the Tm, *e.g.*, 15°C, 20°C, 25°C or 30°C lower than the Tm. The effect of salt concentration (in M of NaCl) can also be calculated, see for example, Brown, A., "Hybridization" pp. 503-506, in *The Encyclopedia of Molec. Biol.*, J. Kendrew, Ed., Blackwell, Oxford (1994).

Preferably, the nucleic acid sequence of a MarA family member identified in this way is at least about 10%, 20%, more preferably at least about 30%, more preferably at least about 40% identical and preferably at least about 50%, or 60% identical to a MarA nucleotide sequence. In preferred embodiments, the nucleic acid sequence of a MarA family member is at least about 70%, 80%, preferably at least about 90%, more preferably at least about 95% identical with a MarA nucleotide sequence. Preferably, MarA family members have an amino acid sequence at least about 20%, preferably at least about 30%, more preferably at least about 40% identical and preferably at least about 50%, or 60% or more identical with a MarA amino acid sequence. In preferred embodiments, the nucleic acid sequence of a MarA family member is at least about 70%, 80%, more preferably at least about 90%, or more preferably at least about 95% identical with a MarA nucleotide sequence. However, it will be understood that the level of sequence similarity among microbial regulators of gene transcription, even though members of the same family, is not necessarily high. This is particularly true in the case of divergent genomes where the level of sequence identity may be low, *e.g.*, less than 20% (*e.g.*, *B. burgdorferi* as compared *e.g.*, to *B.*

subtilis). Accordingly, structural similarity among MarA family members can also be determined based on "three-dimensional correspondence" of amino acid residues. As used herein, the language "three-dimensional correspondence" is meant to includes residues which spatially correspond, *e.g.*, are in the same position of a MarA family polypeptide member as determined, *e.g.*, by x-ray crystallography, but which may not correspond when aligned using a linear alignment program. The language "three-dimensional correspondence" also includes residues which perform the same function, *e.g.*, bind to DNA or bind the same cofactor, as determined, *e.g.*, by mutational analysis.

Exemplary MarA family polypeptides are shown in Table 1, Figures 1-3, and at Prosite (PS00041) and include: AarP, Ada, AdaA, AdiY, AfrR, AggR, AppY, AraC, CafR, CelD, CfaD, CsvR, D90812, EnvY, ExsA, FapR, HrpB, InF, InvF, LcrF, LumQ, MarA, MelR, MixE, MmsR, MsmR, OrfR, Orf_f375, PchR, PerA, PocR, PqrA, RafR, RamA, RhaR, RhaS, Rns, Rob, SoxS, S52856, TetD, TcpN, ThcR, TmbS, U73857, U34257, U21191, UreR, VirF, XylR, XylS, Xys1, 2, 3, 4, Ya52, YbbB, YfiF, YisR, YzbC, and YijO. The nucleotide and amino acid sequences of the *E. coli* Rob molecule are shown in SEQ ID NO:3 and 4, respectively.

TABLE 1. Some Bacterial MarA homologs[a]

Gram-negative bacteria		Gram-positive bacteria
Escherichia coli	*Kiebsiella pneumoniae*	*Lactobacillus helveticus*
MarA (1)	**RamA** (27)	**U34257** (38)
OrfR (2, 3)		
SoxS (4, 5)	*Haemophilus influenzae*	*Azorhizobium caulinodans*
AfrR (6)	Ya52 (28)	S52856 (39)
AraC (7)		
CelD (8)	*Yersinia spp.*	*Streptomyces spp.*
D90812 (9)	CafR (29)	**U21191** (40)
FapR (10, 11)	LcrF (30) or VirF (30)	AraL (41)
MelR (12)		
ORF *f*375 (13, 14)	*Providencia stuartii*	*Streptococcus mutans*
RhaR (15, 16, 17)	**AarP** (31)	MsmR (42)
RhaS (18)		
Rob (19)	*Pseudomonas spp.*	*Pediococcus pentosaceus*
U73857 (20)	MmsR (32)	RafR (43)
XylR (21)	TmbS (33)	
YijO (22)	XylS (34)	*Photobacterium leiognathi*
	Xys1,2,3,4 (35, 36)	LumQ (44)
Proteus vulgaris		
PqrA (23)	Cyanobacteria	*Bacillus subtilis*
	Synechocystis spp.	AdaA (45)
Salmonella typhimurium	**LumQ** (37)	YbbB (46)
MarA (24)	PchR (37)	YfiF (47)
InvF (25)		YisR (48)
PocR (26)		YzbC (49)

[a] The smaller MarA homologs, ranging in size from 87 (U34257) to 138 (OrfR) amino acid residues, are represented in boldface. References are given in parentheses and are listed below.

References for Table 1:

(1) S.P. Cohen, *et al.* 1993. *J. Bacteriol.* 175:1484-1492
(2) G.M. Braus, *et al.* 1984. *J. Bacteriol.* 160:504-509
(3) K. Schollmeier, *et al.*, 1984. *J. Bacteriol.* 160:499-503
(4) C.F. Amabile-Cuevas, *et al.*, 1991. *Nucleic Acids Res.* 19:4479-4484

(5) J. Wu, *et al.*, 1991. *J. Bacteriol.* 173:2864-2871
(6) M.K. Wolf, *et al.*, 1990. *Infect. Immun.* 58:1124-1128
(7) C.M. Stoner, *et al.* 1982. *J. Mol. Biol.* 153:649-652
(8) L.L. Parker, *et al.*, 1990. *Genetics* 123:455-471
(9) H. Mori, 1996. Unpublished data taken from the NCBI databases
(10) P. Klaasen, *et al.*, 1990. *Mol. Microbiol.* 4:1779-1783
(11) M. Ahmed, *et al.*, 1994. *J. Biol. Chem* 269-28506-28513
(12) C. Webster, *et al.*, 1989. *Gene* 83:207-213
(13) G. Plunkett, III. 1995. Unpublished
(14) C Garcia-Martin, *et al.*, 1992. *J. Gen. Microbiol.* 138:1109-1116
(15) G. Plunkett, III., *et al.* 1993. *Nucleic Acids Res.* 21:3391-3398
(16) C. G. Tate, *et al.* 1992. *J. Biol. Chem.* 267:6923-6932
(17) J.F. Tobin *et al.*, 1987. *J. Mol. Biol.* 196:789-799
(18) J. Nishitani, 1991. *Gene* 105:37-42
(19) R.E. Benz, *et al.*, 1993. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt.1 Orig.* 278:187-196
(20) M. Duncan, *et al.*, 1996. Unpublished data
(21) H.J. Sofia, *et al.*, 1994. *Nucleic Acids Res.* 22:2576-2586
(22) F.R. Blattner, *et al.*, 1993. *Nucleic Acids Res.* 21:5408-5417
(23) H. Ishida, *et al.*, 1995. *Antimicrob. Agents Chemother.* 39:453-457
(24) M.C. Sulavik, *et al.*, 1997. *J. Bacteriol.* 179:1857-1866
(25) K. Kaniga, *et al.*, 1994. *Mol. Microbiol.* 13:555-568
(26) J.R. Roth, *et al.* 1993. *J. Bacteriol.* 175:3303-3316
(27) A.M. George, *et al.*, 1983. *J. Bacteriol.* 155:541-548
(28) R.D. Fleischmann, *et al.*, 1995. *Science* 269:469-512
(29) E.E. Galyov, *et al.*, 1991. *FEBS Lett.* 286:79-82
(30) N.P. Hoe, *et al.*, 1992. *J. Bacteriol.* 174:4275-4286
(31) G. Cornelis, *et al.*, 1989. *J. Bacteriol.* 171:254-262
(32) D.R. Macinga, *et al.*, 1995. *J. Bacteriol.* 177:3407-3413
(33) M.I. Steele, *et al.*, 1992. *J. Biol. Chem.* 267:13585-13592
(34) G. Deho, *et al.*, 1995. Unpublished data
(35) N. Mermod, *et al.*, 1984. *EMBO J.* 3:2461-2466
(36) S.J. Assinder, *et al.*, 1992. *Nucleic Acids Res.* 20:5476
(37) S.J. Assinder, *et al.*, 1993. *J. Gen. Microbiol.* 139:557-568
(38) E.G. Dudley, *et al.*, 1996. *J. Bacteriol.* 178:701-704
(39) D. Geelen, *et al.*, 1995. Unpublished data
(40) J. Kormanec, *et al.*, 1995. *Gene* 165:77-80

(41) C.W. Chen, *et al.*, 1992. *J. Bacteriol.* 174:7762-7769
(42) R.R. Russell, *et al.*, 1992. *J. Biol. Chem*, 267:4631-4637
(43) K.K. Leenhouts, *et al.*, 1995. Unpublished data
(44) J.W. Lin, *et al.*, 1995. *Biochem. Biophys. Res. Commun.* 217:684-695
(45) F. Morohoshi, *et al.* 1990. *Nucleic Acids Res.* 18:5473-5480
(46) M. Rosenberg, *et al.*, 1979. *Annu. Rev. Genet.* 13:319-353
(47) H. Yamamoto, *et al.*, 1996. *Microbiology* 142:1417-1421
(48) L.B. Bussey, *et al.*, 1993. *J. Bacteriol.* 175:6348-6353
(49) P.G. Quirk, *et al.*, 1994. *Biochim. Biophys. Acta* 1186:27-34

The term "transcription factor modulating compound" or transcription factor modulator" includes HTH protein modulating compounds, HTH protein modulators. Transcription factor modulating compounds include compounds which interact with one or more transcription factors, such that the activity of the transcription factor is modulated, e.g., enhanced or inhibited. The term also includes both AraC family modulating compounds and MarA family modulating compounds. In one embodiment, the transcription factor modulating compound is an inhibiting compound of a transcription factor, e.g., a prokaryotic transcription factor or a eukaryotic , transcription activation factor. In one embodiment, the transcription factor modulating compounds modulate the activity of a transcription factor as measured by assays known in the art or LANCE assays such as those described in Example 8. In one embodiment, the transcription factor modulating compound inhibits a particular transcription factor by about 10% or greater, about 40% or greater, about 50% or greater, about 60% or greater, about 70% or greater, about 80% or greater, about 90% or greater, about 95% or greater, or about 100% as compared to the activity of the transcription factor with out the transcription factor modulating compound. In another embodiment, the transcription factor modulating compound inhibits biofilm formation. In one embodiment, the transcription factor modulating compound inhibits biofilm formation as measured by assays known in the art or the Crystal Violet assay described in Example 7. In one embodiment, the transcription factor of the invention inhibits the formation of a biofilm by about 25% or more, 50% or more, 75% or more, 80% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.9% or more, 99.99% or more, or by 100%, as compared to the formation of a biofilm without the transcription factor modulating compound.

The term "HTH protein modulating compound" or "HTH protein modulator" includes compounds which interact with one or more HTH proteins such that the activity of the HTH protein is modulated, e.g., enhanced or, inhibited. In one

embodiment, the HTH protein modulating compound is a MarA family polypeptide modulating compound. In one embodiment, the activity of the HTH protein is enhanced when it interacts with the HTH protein modulating compound. For example, the activity of the HTH protein may be increased by greater than 10%, greater the 20%, greater than 50%, greater than 75%, greater than 80%, greater than 90%, or 100% of the activity of the HTH protein in the absence of the HTH modulating compound. In another embodiment, the activity of the HTH protein is decreased upon an interaction with the HTH protein modulating compound. In an embodiment, the activity of the HTH protein is decreased by about 25% or more, 50% or more, 75% or more, 80% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, 99.9% or more, 99.99% or more, or by 100%, as compared to the activity of the protein of a HTH protein when not contacted with an HTH modulating compound of the invention using techniques and assays described herein. Values and ranges included and/or intermediate of the values set forth herein are also intended to be within the scope of the present invention.

The term "MarA family polypeptide modulating compound" or "MarA family modulating compound" include compounds which interact with one or more MarA family polypeptides such that the activity of the MarA family peptide is enhanced or inhibited. In an embodiment, the MarA family polypeptide modulating compound is an inhibiting compound. In a further embodiment, the MarA family inhibiting compound is an inhibitor of MarA, Rob, and/or SoxS. In another embodiment, the MarA family polypeptide modulating compound modulates the expression of luciferase in the Luciferase Assay described in Example 9. In one embodiment, the MarA family polypeptide modulating compound decreases luciferase expression by greater than 10%, greater than 20%, greater than 30%, greater than 40%, greater than 50%, greater than 60%, greater than 70%, greater than 80%, greater than 90% or about 100%.

The term "polypeptide(s)" refers to a peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" includes both short chains, commonly referred to as peptides, oligopeptides and oligomers and longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification may be present in the

same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. Modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins, such as arginylation, and ubiquitination. See, for instance, Proteins--Structure And Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993) and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in Posttranslational Covalent Modification Of Proteins, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990) and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

Preferred polypeptides (and the nucleic acid molecules that encode them) are "naturally occurring." As used herein, a "naturally-occurring" molecule refers to a molecule having an amino acid or a nucleotide sequence that occurs in nature (*e.g.*, a natural polypeptide). In addition, naturally or non-naturally occurring variants of the polypeptides and nucleic acid molecules which retain the same functional activity, (such as, the ability to bind to target nucleic acid molecules (*e.g.*, comprising a marbox) or to polypeptides (*e.g.* RNA polymerase) with a naturally occurring polypeptide are provided for. Such immunologic cross-reactivity can be demonstrated, *e.g.*, by the ability of a variant to bind to a MarA family polypeptide responsive element. Such variants can be made, *e.g.*, by mutation using techniques that are known in the art. Alternatively, variants can be chemically synthesized.

As used herein the term "variant(s)" includes nucleic acid molecules or polypeptides that differ in sequence from a reference nucleic acid molecule or polypeptide, but retain its essential properties. Changes in the nucleotide sequence of the variant may, or may not, alter the amino acid sequence of a polypeptide encoded by the reference nucleic acid molecule. Nucleotide or amino acid changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by a naturally occurring reference sequence. A typical variant of a polypeptide differs in amino acid sequence from a reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, and/or deletions in any combination.

A variant of a nucleic acid molecule or polypeptide may be naturally occurring, such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of nucleic acid molecules and polypeptides may be made from a reference nucleic acid molecule or polypeptide by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans. Alternatively, variants can be chemically synthesized. For instance, artificial or mutant forms of autologous polypeptides which are functionally equivalent, (*e.g.*, have the ability to interact with a MarA family polypeptide responsive element) can be made using techniques which are well known in the art.

Mutations can include, *e.g.*, at least one discrete point mutation which can give rise to a substitution, or by at least one deletion or insertion. For example, mutations can also be made by random mutagenesis or using cassette mutagenesis. For the former, the entire coding region of a molecule is mutagenized by one of several methods (chemical, PCR, doped oligonucleotide synthesis) and that collection of randomly mutated molecules is subjected to selection or screening procedures. In the latter, discrete regions of a polypeptide, corresponding either to defined structural or functional determinants are subjected to saturating or semi-random mutagenesis and these mutagenized cassettes are re-introduced into the context of the otherwise wild type allele. In one embodiment, PCR mutagenesis can be used. For example, Megaprimer PCR can be used (O.H. Landt, 1990. Gene 96:125-128).

In preferred embodiments, a MarA family polypeptide excludes one or more of XylS, AraC, and MelR. In other preferred embodiments, a MarA family polypeptide is involved in antibiotic resistance. In particularly preferred embodiments, a MarA family

polypeptide is selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.

The language "activity of a transcription factor" includes the ability of a transcription factor to interact with DNA, *e.g.*, to bind to a transcription factor responsive promoter, or to initiate transcription from such a promoter. The language expressly includes the activities of AraC family polypeptides, HTH proteins and MarA family polypeptides.

The language "activity of a MarA family polypeptide" includes the ability of the MarA family polypeptide to interact with DNA, *e.g.*, to bind to a MarA family polypeptide responsive promoter, or to initiate transcription from such a promoter. MarA functions both as a transcriptional activator (*e.g.*, upregulating genes such as inaA, galT, micF, etc.) and as a repressor (*e.g.*, downregulating genes such as fecA, purA, guaB, etc.) (Alekshun, 1997, *Antimicrob. Agents Chemother.* 41:2067-2075; Barbosa & Levy, *J. Bact.* 2000, Vol. 182, p. 3467-3474; Pomposiello *et al. J. Bact.* 2001, Vol 183, p. 3890-3902).

The language "transcription factor responsive element" includes a nucleic acid sequence which can interact with promoters or enhancers which are involved in initiating transcription of an operon in a microbe. The language includes marA family polypeptide responsive elements.

The language "marA family polypeptide responsive element" includes a nucleic acid sequence which can interact with marA, *e.g.*, promoters or enhancers which are involved in regulating transcription of a nucleic acid sequence in a microbe. MarA responsive elements comprise approximately 16 base pair marbox sequence, the sequence critical for the binding of MarA to its target. In addition, a secondary site, the accessory marbox, upstream of the primary marbox contributes to basal and derepressed mar transcription. A marbox may be situated in either the forward or backward orientation. (Martin, 1999, *Mol. Microbiol.* 34:431-441). In the *marRAB* operon, the marbox is in the backward orientation and is thus located on the sense strand with respect to *marRAB* (Martin, 1999, *Mol. Microbiol.* 34:431-441). Subtle differences within the marbox sequence of particular promoters may account for differential regulation by MarA and other related, *e.g.*, SoxS and Rob, transcription factors (Martin, 2000, *Mol Microbiol*; 35(3):623-34). In one embodiment, MarA family responsive elements are promoters that are structurally or functionally related to a *marA* promoter, *e.g.*, interact with MarA or a protein related to MarA. Preferably, the *marA* family polypeptide responsive element is a *marRAB* promoter. For example, in the *mar* operon, several promoters are *marA* family polypeptide responsive promoters as defined herein, *e.g.*, the 405-bp *ThaI* fragment from the *marO* region is a *marA* family responsive

promoter (Cohen *et al.* 1993. *J. Bact.* 175:7856). In addition, MarA has been shown to bind to a 16 bp MarA binding site (referred to as the "marbox" within *marO* (Martin *et al.* 1996. *J. Bacteriol.* 178:2216). MarA also affects transcription from the *acrAB; micF; mlr 1,2,3; slp; nfo; inaA; fpr; sodA; soi-17,19; zwf; fumC; or rpsF* promoters (Alekshun and Levy. 1997. *Antimicrobial Agents and Chemother.* 41:2067). Other *marA* family responsive promoters are known in the art and include: *araBAD, araE, araFGH and araC*, which are activated by AraC; Pm, which is activated by XylS; *melAB* which is activated by MelR; and *oriC* which is bound by Rob.

The language "MarA family polypeptide responsive promoter" also includes portions of the above promoters which are sufficient to activate transcription upon interaction with a MarA family member protein. The portions of any of the MarA family polypeptide-responsive promoters which are minimally required for their activity can be easily determined by one of ordinary skill in the art, e.g., using mutagenesis. Exemplary techniques are described by Gallegos *et al.* (1996, *J. Bacteriol.* 178:6427). A "MarA family polypeptide responsive promoter" also includes non-naturally occurring variants of MarA family polypeptide responsive promoters which have the same function as naturally occurring MarA family promoters. Preferably such variants have at least 30% or greater, 40% or greater, or 50% or greater, nucleotide sequence identity with a naturally occurring MarA family polypeptide responsive promoter. In preferred embodiments, such variants have at least about 70% nucleotide sequence identity with a naturally occurring MarA family polypeptide responsive promoter. In more preferred embodiments, such variants have at least about 80% nucleotide sequence identity with a naturally occurring MarA family polypeptide responsive promoter. In particularly preferred embodiments, such variants have at least about 90% nucleotide sequence identity and preferably at least about 95% nucleotide sequence identity with a naturally occurring MarA family polypeptide responsive promoter. In yet other embodiments nucleic acid molecules encoding variants of MarA family polypeptide responsive promoters are capable of hybridizing under stringent conditions to nucleic acid molecules encoding naturally occurring MarA family polypeptide responsive promoters.

The term "interact" includes close contact between molecules that results in a measurable effect, *e.g.*, the binding of one molecule with another. For example, a MarA family polypeptide can interact with a MarA family polypeptide responsive element and alter the level of transcription of DNA. Likewise, compounds can interact with a MarA family polypeptide and alter the activity of a MarA family polypeptide.

The term "inducible promoter" includes promoters that are activated to induce the synthesis of the genes they control. As used herein, the term "constitutive promoter" includes promoters that do not require the presence of an inducer, *e.g.*, are continuously active.

The terms "heterologous DNA" or "heterologous nucleic acid" includes DNA that does not occur naturally in the cell (*e.g.*, as part of the genome) in which it is present or which is found in a location or locations in the genome that differ from that in which it occurs in nature or which is operatively linked to DNA to which it is not normally linked in nature (i.e., a gene that has been operatively linked to a heterologous promoter). Heterologous DNA is 1) not naturally occurring in a particular position (*e.g.*, at a particular position in the genome) or 2) is not endogenous to the cell into which it is introduced, but has been obtained from another cell. Heterologous DNA can be from the same species or from a different species. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which is expressed is herein encompassed by the term heterologous DNA.

The terms "heterologous protein", "recombinant protein", and "exogenous protein" are used interchangeably throughout the specification and refer to a polypeptide which is produced by recombinant DNA techniques, wherein generally, DNA encoding the polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. That is, the polypeptide is expressed from a heterologous nucleic acid molecule.

The term "microbe" includes microorganisms expressing or made to express a transcription factor, araC family polypeptide, HTH protein, or a marA family polypeptide. "Microbes" are of some economic importance, *e.g.*, are environmentally important or are important as human pathogens. For example, in one embodiment microbes cause environmental problems, *e.g.*, fouling or spoilage, or perform useful functions such as breakdown of plant matter. In another embodiment, microbes are organisms that live in or on mammals and are medically important. Preferably microbes are unicellular and include bacteria, fungi, or protozoa. In another embodiment, microbes suitable for use in the invention are multicellular, *e.g.*, parasites or fungi. In preferred embodiments, microbes are pathogenic for humans, animals, or plants. Microbes may be used as intact cells or as sources of materials for cell-free assays. In one embodiment, the microbes include prokaryotic organisms. In other embodiments, the microbes include eukaryotic organisms.

The term selective marker includes polypeptides that serve as indicators, *e.g.*, provide a selectable or screenable trait when expressed by a cell. The term "selective marker" includes both selectable markers and counterselectable markers. As

used herein the term "selectable marker" includes markers that result in a growth advantage when a compound or molecule that fulfills the test parameter of the assay is present. The term "counterselectable marker" includes markers that result in a growth disadvantage unless a compound or molecule is present which disrupts a condition giving rise to expression of the counterselectable marker. Exemplary selective markers include cytotoxic gene products, gene products that confer antibiotic resistance, gene products that are essential for growth, gene products that confer a selective growth disadvantage when expressed in the presence of a particular metabolic substrate (*e.g.*, the expression of the URA3 gene confers a growth disadvantage in the presence of 5-fluoroorotic acid).

The term "test compound" includes any reagent or test agent which is employed in the assays of the invention and assayed for its ability to influence the activity of a transcription factor, e.g., an AraC family polypeptide, an HTH protein, or a MarA family polypeptide, *e.g.*, by binding to the polypeptide or to a molecule with which it interacts. More than one compound, *e.g.*, a plurality of compounds, can be tested at the same time for their ability to modulate the activity of a transcription factor, e.g., an AraC family polypeptide, an HTH protein, or a MarA family polypeptide, activity in a screening assay. In an advantageous embodiment, the test compound is a MarA family modulating compound.

Test compounds that can be tested in the subject assays include antibiotic and non-antibiotic compounds. In one embodiment, test compounds include candidate detergent or disinfectant compounds. Exemplary test compounds which can be screened for activity include, but are not limited to, peptides, non-peptidic compounds, nucleic acids, carbohydrates, small organic molecules (*e.g.*, polyketides), and natural product extract libraries. The term "non-peptidic test compound" includes compounds that are comprised, at least in part, of molecular structures different from naturally-occurring L-amino acid residues linked by natural peptide bonds. However, "non-peptidic test compounds" also include compounds composed, in whole or in part, of peptidomimetic structures, such as D-amino acids, non-naturally-occurring L-amino acids, modified peptide backbones and the like, as well as compounds that are composed, in whole or in part, of molecular structures unrelated to naturally-occurring L-amino acid residues linked by natural peptide bonds. "Non-peptidic test compounds" also are intended to include natural products.

The term "antagonist" includes transcription factor modulating compounds (e.g., AraC family polypeptide modulating compounds, HTH protein modulating compounds, MarA family polypeptide modulating compounds, etc.) which inhibit the activity of a transcription factor by binding to and inactivating the

transcription factor (e.g., an AraC family modulating compound, an MarA family polypeptide modulating compound, etc.), by binding to a nucleic acid target with which the transcription factor interacts (*e.g.*, for MarA, a marbox), by disrupting a signal transduction pathway responsible for activation of a particular regulon (*e.g.*, for Mar, the inactivation of MarR or activation of MarA synthesis), and/or by disrupting a critical protein-protein interaction (*e.g.*, MarA-RNA polymerase interactions that are required for MarA to function as a transcription factor.) Antagonists may include, for example, naturally or chemically synthesized compounds such as small cell permeable organic molecules, nucleic acid interchelators, peptides, *etc.*

The term "agonist" includes transcription factor modulating compounds (e.g., AraC family polypeptide modulating compounds, HTH protein modulating compounds, MarA family polypeptide modulating compounds, etc.) which promote the activity of a transcription factor by binding to and activating the transcription factor (e.g., an AraC family modulating compound, an MarA family polypeptide modulating compound, etc.), by binding to a nucleic acid target with which the transcription factor interacts (*e.g.*, for MarA, a marbox), by facilitating a signal transduction pathway responsible for activation of a particular regulon (*e.g.*, for Mar, the inactivation of MarR or activation of MarA synthesis), and/or by facilitating a critical protein-protein interaction (*e.g.*, MarA-RNA polymerase interactions that are required for MarA to function as a transcription factor.) Agonists may include, for example, naturally or chemically synthesized compounds such as small cell permeable organic molecules, nucleic acid interchelators, peptides, *etc.*

II. MarA Family polypeptide Helix-Turn-Helix Domains

Helix-turn-helix domains are known in the art and have been implicated in DNA binding (*Ann Rev. of Biochem.* 1984. 53:293). An example of the consensus sequence for a helix-turn domain can be found in Brunelle and Schleif (1989, *J. Mol. Biol.* 209:607). The domain has been illustrated by the sequence XXXPhoAlaXXPhoGlyPhoXXXXPhoXXPhoXX, where X is any amino acid and Pho is a hydrophobic amino acid.

The crystal structure of MarA has been determined and the first (most amino terminal) HTH domain of MarA has been identified as comprising from about amino acid 31 to about amino acid 52 and the second HTH domain of MarA has been identified as comprising from about amino acid 79 to about amino acid 102 (Rhee et al. 1998. *Proc. Natl. Acad. Sci. USA*. 95:10413).

Locations of the helix-turn-helix domains in other MarA family members as well as other HTH proteins can easily be found by one of skill in the art. For example using the MarA protein sequence and an alignment program, *e.g.*, the ProDom program or other programs known in the art, a portion of the MarA amino acid sequence *e.g.*, comprising one or both HTH domains of MarA (such as from about amino acid 30 to about amino acid 107 of MarA) to produce an alignment. Using such an alignment, the amino acid sequences corresponding to the HTH domains of MarA can be identified in other MarA family member proteins. An exemplary consensus sequence for the first helix-turn-helix domain of a MarA family polypeptide can be illustrated as XXXXAXXXXXSXXXLXXXFX, where X is any amino acid. An exemplary consensus sequence for the second helix-turn-helix domain of a MarA family polypeptide is illustrated as XXIXXIAXXXGFXSXXXFXXX[F/Y], where X is any amino acid. Preferably, a MarA family polypeptide first helix-turn-helix domain comprises the consensus sequence E/D-X-V/L-A-D/E-X-A/S-G-X-S-X3-L-Q-X2-F-K/R/E-X2-T/I. Preferably, a MarA family polypeptide second helix-turn-helix domain comprises the consensus sequence I-X-D-I-A-X3-G-F-X-S-X2-F-X3-F-X4.

In an embodiment, a MarA family member HTH domain is a MarA HTH domain. The first and second helix-turn-helix domains of MarA are, respectively, EKVSERSGYSKWHLQRMFKKET and ILYLAERYGFESQQTLTRTFKNYF. Other exemplary MarA family helix-turn-helix domains include: about amino acid 210 to about amino acid 229 and about amino acid 259 to about amino acid 278 of MelR; about amino acid 196 to about amino acid 215 and about amino acid 245 to about amino acid 264 of AraC; and about amino acid 230 to about amino acid 249 (or 233-253) and about amino acid 281 to about amino acid 301 (or 282-302) of XylS (see *e.g.*, Brunelle *et al.* 1989. *J. Mol. Biol.* 209:607; Niland *et al.* 1996. *J. Mol. Biol.* 264:667; Gallegos *et al.* 1997. *Microbiology and Molecular Biology Reviews.* 61:393).

"MarA family polypeptide helix-turn-helix domains" are derived from or are homologous to the helix-turn-helix domains found in the MarA family polypeptides as described supra. In preferred embodiments, a MarA family polypeptide excludes one or more of XylS, AraC, and MelR. In particularly preferred embodiments, a MarA family polypeptide is selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.

Both of the helix-turn-helix domains present in MarA family polypeptides are in the carboxy terminal end of the protein. Proteins or portions thereof comprising either or both of these domains can be used in the instant methods. In certain embodiments, a polypeptide which is used in screening for compounds comprises the helix-turn-helix domain most proximal to the carboxy terminus (HTH2)

of the MarA family polypeptide from which it is derived. In other embodiments, such a polypeptide comprises the helix-turn-helix domain most proximal to the amino terminus (HTH1) of the MarA family polypeptide from which it is derived. In one embodiment, other polypeptide sequences may also be present, *e.g.*, sequences that might facilitate immobilizing the domain on a support, or, alternatively, might facilitate the purification of the domain.

In an embodiment, such a polypeptide consists essentially of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family polypeptide from which it is derived. In other preferred embodiments, such a polypeptide consists essentially of the helix-turn-helix domain most proximal to the amino terminus of the MarA family polypeptide from which it is derived.

In an embodiment, such a polypeptide consists of the helix-turn-helix domain most proximal to the carboxy terminus of the AraC family polypeptide or MarA family polypeptide from which it is derived. In other preferred embodiments, such a polypeptide consists of the helix-turn-helix domain most proximal to the amino terminus of the AraC family polypeptide or MarA family polypeptide from which it is derived.

MarA family polypeptide or AraC family polypeptide helix-turn-helix domains can be made using techniques which are known in the art. The nucleic acid and amino acid sequences of transcription factors, such as MarA family polypeptides, are available, for example, from GenBank. Using this information, the helix-turn-helix consensus motif and mutational analysis provided herein, one of ordinary skill in the art can identify MarA family or AraC family polypeptide helix-turn-helix domains.

In certain embodiments of the invention it will be desirable to obtain "isolated or recombinant" nucleic acid molecules encoding transcription factors or portions thereof (e.g., HTH protein helix-turn-helix domains, AraC family helix-turn-helix domains, MarA family helix-turn-helix domains or mutant forms thereof). By "isolated or recombinant" is meant a nucleic acid molecule which has been (1) amplified in vitro by, for example, polymerase chain reaction (PCR); (2) recombinantly produced by cloning, or (3) purified, as by cleavage and gel separation; or (4) synthesized by, for example, chemical synthesis. Such a nucleic acid molecule is isolated from the sequences which naturally flank it in the genome and from cellular components.

The isolated or recombinant nucleic acid molecules encoding transcription factors (e.g., HTH protein helix-turn-helix domains, AraC family helix-turn-helix domains, MarA family helix-turn-helix domains or mutant forms thereof) can then, for example, be utilized in binding assays, can be expressed in a cell, or can be expressed on the surface of phage, as discussed further below.

In yet other embodiments of the invention, it will be desirable to obtain a substantially purified or recombinant HTH protein helix-turn-helix domains (e.g., MarA family helix-turn-helix domains or mutant forms thereof). Such polypeptides, for example, can be purified from cells which have been engineered to express an isolated or recombinant nucleic acid molecule which encodes a HTH protein helix-turn-helix domain (e.g., MarA family helix-turn-helix domain or mutant forms thereof). For example, as described in more detail below, a bacterial cell can be transformed with a plasmid which encodes a MarA family helix-turn-helix domain. The MarA family helix-turn-helix protein can then be purified from the bacterial cells and used, for example, in the cell-free assays described herein.

Purification of a HTH protein helix-turn-helix domain (e.g., MarA family helix-turn-helix domain) can be accomplished using techniques known in the art. For example, column chromatography could be used, or antibodies specific for the domain or for a polypeptide fused to the domain can be employed, for example on a column or in a panning assay.

In preferred embodiments, cells used to express HTH protein helix-turn-helix domains (e.g., MarA family helix-turn-helix domains or mutant forms thereof) for purification, *e.g.*, host cells, comprise a mutation which renders any endogenous HTH proteins nonfunctional or causes the endogenous protein to not be expressed. In other embodiments, mutations may also be made in MarR or related genes of the host cell, such that repressor proteins which bind to the same promoter as a MarA family polypeptide are not expressed by the host cell.

In certain embodiments of the invention, it will be desirable to use a mutant form of a HTH protein helix-turn helix domain, e.g., a non-naturally occurring form of a MarA family helix-turn-helix domain which has altered activity, *e.g.*, does not retain wild type MarA family polypeptide helix-turn-helix domain activity, or which has reduced activity or which is more active when compared to a wild-type MarA family polypeptide helix-turn-helix domain.

Such mutant forms can be made using techniques which are well known in the art. For example, random mutagenesis can be used. Using random mutagenesis one can mutagenize an entire molecule or one can proceed by cassette mutagenesis. In the former instance, the entire coding region of a molecule is mutagenized by one of several methods (chemical, PCR, doped oligonucleotide synthesis) and that collection of randomly mutated molecules is subjected to selection or screening procedures. In the second approach, discrete regions of a protein, corresponding either to defined structural or functional determinants (*e.g.*, the first or second alpha helix of a helix-turn-helix

domain) are subjected to saturating or semi-random mutagenesis and these mutagenized cassettes are re-introduced into the context of the otherwise wild type allele.

In a preferred embodiment, PCR mutagenesis is used. For example, Example 2 describes the use of Megaprimer PCR (O.H. Landt, *Gene* 96:125-128) used to introduce an *Nhe*I restriction site into the centers of both the helix A (position 1989) and helix B (position 2016) regions of the *marA* gene.

In one embodiment, such mutant helix-turn-helix domains comprise one or more mutations in the helix-turn-helix domain most proximal to the carboxy terminus (HTH2) of the MarA family polypeptide molecule. In a preferred embodiment, the mutation comprises an insertion into helix A and helix B of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family polypeptide. In one embodiment, such mutant helix-turn-helix domains comprise one or more mutations in the helix-turn-helix domain most proximal to the amino terminus (HTH1) of the MarA family polypeptide molecule. In a preferred embodiment, the mutation comprises an insertion into helix A and helix B of the helix-turn-helix domain most proximal to the amino terminus of the MarA family polypeptide. In particularly preferred embodiments, the mutation is selected from the group consisting of: an insertion at an amino acid corresponding to about position 33 of MarA and an insertion at an amino acid position corresponding to about position 42 of MarA. "Corresponding" amino acids can be determined, e.g., using an alignment of the helix-turn-helix domains.

Such mutant forms of MarA family helix-turn-helix motifs are useful as controls to verify the specificity of antiinfective compounds for a MarA family helix-turn-helix domain or as controls for the identification of genetic loci which affect resistance to antiinfectives. For example, the mutant MarA family helix-turn-helix domains described in the appended Examples demonstrate that insertional inactivation of MarA at either helix A or helix B in the first HTH domain abolished the multidrug resistance phenotype in both *E. coli* and *M. smegmatis*. By the use of an assay system such as that described in Example 2, which demonstrates the ability of MarA family polypeptide helix-turn-helix domains to increase antibiotic resistance and that mutant forms of these domains do not have the same effect, one can clearly show that the response of any genetic loci identified is specific to a MarA family helix-turn-helix domain.

III. Expression of Polypeptide or Portions Thereof

Nucleic acids encoding transcription factors, such as AraC family polypeptides, HTH proteins, e.g., MarA family polypeptides or selectable markers (or portions thereof that retain an activity of the full-length polypeptide, *e.g.*, are capable of

binding to a transcription factor responsive element or retain their indicator function) can be expressed in cells using vectors. Almost any conventional delivery vector can be used. Such vectors are widely available commercially and it is within the knowledge and discretion of one of ordinary skill in the art to choose a vector which is appropriate for use with a given microbial cell. The sequences encoding these domains can be introduced into a cell on a self-replicating vector or may be introduced into the chromosome of a microbe using homologous recombination or by an insertion element such as a transposon.

These nucleic acids can be introduced into microbial cells using standard techniques, for example, by transformation using calcium chloride or electroporation. Such techniques for the introduction of DNA into microbes are well known in the art. In one embodiment, a nucleic acid molecule which has been amplified *in vitro* by, for example, polymerase chain reaction (PCR); recombinantly produced by cloning, or) purified, as by cleavage and gel separation; or synthesized by, for example, chemical synthesis can be used to produce MarA family polypeptides (George, A. M. & Levy, S. B. (1983)J. *Bacteriol.* 155, 541-548; Cohen, S. P. *et al.* (1993) J *Infect. Dis.* 168, 484-488; Cohen, S. P *et al.* (1993) J *Bacteriol.* 175, 1484-1492; Sulavick, M. C. *et al.* (1997) J. *Bacteriol.* 179, 1857-1866).

Host cells can be genetically engineered to incorporate nucleic acid molecules of the invention. In one embodiment nucleic acid molecules specifying transcription factors can be placed in a vector. The term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid molecule to which it has been linked. The term "expression vector" or "expression system" includes any vector, (*e.g.*, a plasmid, cosmid or phage chromosome) containing a gene construct in a form suitable for expression by a cell (*e.g.*, linked to a promoter). In the present specification, "plasmid" and "vector" are used interchangeably, as a plasmid is a commonly used form of vector. Moreover, the invention is intended to include other vectors which serve equivalent functions. A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

Appropriate vectors are widely available commercially and it is within the knowledge and discretion of one of ordinary skill in the art to choose a vector which is appropriate for use with a given host cell. The sequences encoding a transcription factor, such as, for example, MarA family polypeptides, can be introduced into a cell on a self-replicating vector or may be introduced into the chromosome of a microbe using homologous recombination or by an insertion element such as a transposon.

The expression system constructs may contain control regions that regulate expression. "Transcriptional regulatory sequence" is a generic term to refer to DNA sequences, such as initiation signals, enhancers, operators, and promoters, which induce or control transcription of polypeptide coding sequences with which they are operably linked. It will also be understood that a recombinant gene encoding a transcription factor gene, e.g., an HTH protein gene or an AraC family polypeptide, e.g., MarA family polypeptide, can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring transcription factor gene. Exemplary regulatory sequences are described in Goeddel; *Gene Expression Technology*: *Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). For instance, any of a wide variety of expression control sequences, that control the expression of a DNA sequence when operatively linked to it, may be used in these vectors to express DNA sequences encoding the polypeptide.

Generally, any system or vector suitable to maintain, propagate or express nucleic acid molecules and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., Molecular Cloning, A Laboratory Manual, (supra).

Exemplary expression vectors for expression of a gene encoding a polypeptide and capable of replication in a bacterium, *e.g.*, a gram positive, gram negative, or in a cell of a simple eukaryotic fungus such as a *Saccharomyces* or, *Pichia*, or in a cell of a eukaryotic organism such as an insect, a bird, a mammal, or a plant, are known in the art. Such vectors may carry functional replication-specifying sequences (replicons) both for a host for expression, for example a *Streptomyces*, and for a host, for example, *E. coli*, for genetic manipulations and vector construction. See, *e.g.*, U.S. 4,745,056. Suitable vectors for a variety of organisms are described in Ausubel, F. *et al.*, *Short Protocols in Molecular Biology*, Wiley, New York (1995), and for example, for *Pichia*, can be obtained from Invitrogen (Carlsbad, CA).

Useful expression control sequences, include, for example, the early and late promoters of SV40, adenovirus or cytomegalovirus immediate early promoter, the *lac* system, the *trp* system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage lambda , the control regions for fd coat polypeptide, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, *e.g.*, Pho5, the promoters of the yeast α-mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. A useful translational enhancer sequence is described in U.S. 4,820,639.

In one embodiment, an inducible promoter will be employed to express a polypeptide of the invention. For example, in one embodiment, trp (induced by tryptophan), tac (induced by lactose), or tet (induced by tetracycline) can be used in bacterial cells, or GAL1 (induced by galactose) can be used in yeast cell.

In another embodiment, a constitutive promoter can be used to express a polypeptide of the invention.

It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of polypeptide desired to be expressed. Representative examples of appropriate hosts include bacterial cells, such as gram positive, gram negative cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2 and Spodoplera Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

In one embodiment, cells used to express heterologous polypeptides of the invention, comprise a mutation which renders one or more endogenous transcription factors, such as a AraC family polypeptide or a MarA family polypeptide, nonfunctional or causes one or more endogenous polypeptide to not be expressed. Manipulation of the genetic background in this manner allows for screening for compounds that modulate specific transcription factors, such as MarA family members or AraC family members, or more than one transcription factors.

In other embodiments, mutations may also be made in other related genes of the host cell, such that there will be no interference from the endogenous host loci. In yet another embodiment, a mutation may be made in a chromosomal gene to create a heterotroph.

Introduction of a nucleic acid molecule into the host cell ("transformation") can be effected by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology, (1986) and Sambrook et al., Molecular

Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Examples include calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Purification of polypeptides, *e.g.*, recombinantly expressed polypeptides, can be accomplished using techniques known in the art. For example, if the polypeptide is expressed in a form that is secreted from cells, the medium can be collected. Alternatively, if the polypeptide is expressed in a form that is retained by cells, the host cells can be lysed to release the polypeptide. Such spent medium or cell lysate can be used to concentrate and purify the polypeptide. For example, the medium or lysate can be passed over a column, *e.g.*, a column to which antibodies specific for the polypeptide have been bound. Alternatively, such antibodies can be specific for a second polypeptide which has been fused to the first polypeptide (*e.g.*, as a tag) to facilitate purification of the first polypeptide. Other means of purifying polypeptides are known in the art.

IV. Methods of Identifying Antimicrobial/Antiinfective Compounds Which Modulate an Activity of a Transcription factor

In one embodiment, the invention provides for methods of identifying a test compound which modulates the activity of a transcription factor, (e.g., an HTH protein, a MarA family polypeptide, an AraC family polypeptide, etc.) by contacting a cell expressing a transcription factor (or portion thereof) with a test compound under conditions which allow interaction of the test compound with the cell. The ability of the test compound to modulate an activity of a transcription factor can be determined in a variety of ways, as outlined in more detail below. Compounds identified using the subject methods are useful, *e.g.*, to interfere with the ability of a microbe to grow, *e.g.*, on surfaces or in a host or cause infection in a host.

Assays

In one embodiment, the expression of a selectable marker that confers a selective growth disadvantage or lethality is placed under the direct control of a MarA responsive element in a cell expressing marA.

In one embodiment, marA is plasmid encoded. In one embodiment, the genetic background of the host organism is manipulated, *e.g.*, to delete one or more chromosomal marA genes or marA homolog genes.

In one embodiment, expression of marA is controlled by a highly regulated and inducible promoter. For example, in one embodiment, a promoter selected from the group consisting of *trp*, *tac*, or *tet* in bacterial cells or *GAL1* in yeast cells can be used.

In another embodiment, expression of marA is constitutive.

In one embodiment, a selective marker is a cytotoxic gene product (e.g., ccdB).

In another embodiment, a selective marker is a gene that confers antibiotic resistance (*e.g.*, kan, cat, or bla).

In another embodiment, a selective marker is an essential gene (*e.g.*, purA or guaB in a purine or guanine heterotroph).

In still another embodiment, a selective marker is a gene that confers a selective growth disadvantage in the presence of a particular metabolic substrate (*e.g.*, the expression of URA3 in the presence of 5-fluoroorotic acid [5-FOA] in yeast).

In one embodiment, compounds that modulate transcription factors (e.g., HTH proteins, AraC family polypeptides, or MarA family polypeptides) are identified using a one-hybrid screening assay. As used herein, the term "one-hybrid screen" as used herein includes assays that detect the disruption of protein-nucleic acid interactions. These assays will identify agents that interfere with the binding of a transcription factor (e.g., an HTH protein, a AraC family polypeptide, or a MarA family polypeptide) to a particular target, *e.g.*, DNA containing, for MarA, a marbox, at the level of the target itself, *e.g.*, by binding to the target and preventing the trnscriptional activation factor from interacting with or binding to this site.

In another embodiment, compounds of the invention are identified using a two-hybrid screening assay. As used herein the term "two-hybrid screen" as used herein includes assays that detect the disruption of protein-protein interactions. Such two hybrid assays can be used to interfere with crucial protein-transcription factor interactions (e.g., HTH protein interactions, AraC family polypeptide interactions, MarA family polypeptide interactions). One example would be to prevent RNA polymerase-MarA family polypeptide contacts, that are necessary for the MarA family polypeptide to function as a transcription factor (either positive acting or negative acting).

In yet another embodiment, compounds of the invention are identified using a three-hybrid screening assay. As used herein the term "three-hybrid screen" as used herein includes assays that will detect the disruption of a signal transduction pathway(s) required for the activation of a particular regulon of interest. In one embodiment, the three-hybrid screen is used to detect disruption of a signal transduction pathway(s) required for the activation of the Mar regulon, i.e., synthesis of MarA. (Li

and Park. *J. Bact.* 181:4824). The assay can be used to identify compounds that may be responsible for activating transcription factor expression, *e.g.*, Mar induction by antibiotics may proceed in this manner.

In one embodiment of the assay, the expression of a selective marker (*e.g.*, ccdB, cat, bla, kan, guaB, URA3) is put under the direct control of an activatable MarA responsive activatable promoter (*e.g.*, inaA, galT, micF). In the absence of Mar A, the expression of the selective marker would be silent. For example, in the case of regulation of the cytotoxic gene ccdB, the gene would be silent and the cells would survive. Synthesis of MarA from an inducible plasmid in a suitable host would result in the activation of the MarA responsive activatable promoter and expression of the selective marker. In the case of ccdB, the gene would be expressed and result in cell death. Compounds that inhibit MarA would be identified as those that permit cell survival under conditions of MarA expression.

In another embodiment, *e.g.*, where the expression of the MarA responsive activatable promoter regulates a gene such as URA3, a different result could be obtained. In this case, in the absence of MarA and thus, in the absence of URA3 expression, cells would grow in the presence of a 5-FOA. Upon activation of MarA expression and thus synthesis of URA3, cells would die following the conversion of 5-FOA to a toxic metabolite by URA3.

In another embodiment, a selectable marker is put under the direct control of a repressible MarA responsive promoter (*e.g.*, fecA). In this example, under conditions of constitutive MarA synthesis, *e.g.*, in a constitutive mar (marc) mutant the expression of the selectable marker would be silent. In the case of ccdB, this would mean that cells would remain viable. Following inactivation of MarA, the selectable marker would be turned on, resulting in cell death.

In another embodiment, a purine or guanine heterotroph can be constructed by the inactivation of the chromosomal guaB or purA genes in *E. coli*. The guaB or purA gene would then be cloned into a suitable vector, under the control of its natural promoter. This construct would then be transformed into the heterotrophic host. The heterotroph will not grow if MarA expression is constitutive and if cells are grown on media lacking purines or guanine. This can be attributed to MarA mediated repression of guaB or purA synthesis. Candidate inhibiting compounds of MarA can be identified as compounds that restored growth, i.e., relieved MarA mediated repression of guaB and purA expression. In another embodiment, genes that are required for growth *in vivo*, for example in an animal model of infection.

In preferred embodiments, controls may be included to ensure that any compounds which are identified using the subject assays do not merely appear to modulate the activity of a transcription factor, because they inhibit protein synthesis. For example, if a compound appears to inhibit the synthesis of a protein being translated from RNA which is transcribed upon activation of a MarA family responsive element, it may be desirable to show that the synthesis of a control, *e.g.*, a protein which is being translated from RNA which is not transcribed upon activation of a MarA family responsive element, is not affected by the addition of the same compound. For example, the amount of the MarA family polypeptide being made and compared to the amount of an endogenous protein being made. In another embodiment the microbe could be transformed with another plasmid comprising a promoter which is not a MarA family responsive promoter and a protein operably linked to that promoter. The expression of the control protein could be used to normalize the amount of protein produced in the presence and absence of compound.

V. Microbes Suitable For Testing

Numerous different microbes are suitable for testing in the instant assays. As such, they may be used as intact cells or as sources of material, *e.g.*, nucleic acid molecules or polypeptides as described herein.

In preferred embodiments, microbes for use in the claimed methods are bacteria, either Gram negative or Gram positive bacteria. More specifically, any bacteria that are shown to become resistant to antibiotics , *e.g.*, to display a Mar phenotype are preferred for use in the claimed methods, or that are infectious or potentially infectious.

Examples of microbes suitable for testing include, but are not limited to, *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, *Pseudomonas acidovorans*, *Pseudomonas alcaligenes*, *Pseudomonas putida*, *Stenotrophomonas maltophilia*, *Burkholderia cepacia*, *Aeromonas hydrophilia*, *Escherichia coli*, *Citrobacter freundii*, *Salmonella typhimurium*, *Salmonella typhi*, *Salmonella paratyphi*, *Salmonella enteritidis*, *Shigella dysenteriae*, *Shigella flexneri*, *Shigella sonnei*, *Enterobacter cloacae*, *Enterobacter aerogenes*, *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Serratia marcescens*, *Francisella tularensis*, *Morganella morganii*, *Proteus mirabilis*, *Proteus vulgaris*, *Providencia alcalifaciens*, *Providencia rettgeri*, *Providencia stuartii*, *Acinetobacter calcoaceticus*, *Acinetobacter haemolyticus*, *Yersinia enterocolitica*, *Yersinia pestis*, *Yersinia pseudotuberculosis*, *Yersinia intermedia*, *Bordetella pertussis*, *Bordetella parapertussis*, *Bordetella bronchiseptica*, *Haemophilus influenzae*, *Haemophilus parainfluenzae*, *Haemophilus haemolyticus*, *Haemophilus*

parahaemolyticus, *Haemophilus ducreyi*, *Pasteurella multocida*, *Pasteurella haemolytica*, *Branhamella catarrhalis*, *Helicobacter pylori*, *Campylobacter fetus*, *Campylobacter jejuni*, *Campylobacter coli*, *Borrelia burgdorferi*, *Vibrio cholerae*, *Yibrio parahaemolyticus*, *Legionella pneumophila*, *Listeria monocytogenes*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Gardnerella vaginalis*, *Bacteroides fragilis*, *Bacteroides distasonis*, *Bacteroides* 3452A homology group, *Bacteroides vulgatus*, *Bacteroides ovalus*, *Bacteroides thetaiotaomicron*, *Bacteroides uniformis*, *Bacteroides eggerthii*, *Bacteroides splanchnicus*, *Clostridium difficile*, *Mycobacterium tuberculosis*, *Mycobacterium avium*, *Mycobacterium intracellulare*, *Mycobacterium leprae*, *Corynebacterium diphtheriae*, *Corynebacterium ulcerans*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Enterococcus faecalis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus saprophyticus*, *Staphylococcus intermedius*, *Staphylococcus hyicus subsp. hyicus*, *Staphylococcus haemolyticus*, *Staphylococcus hominis*, and *Staphylococcus saccharolyticus*.

In one embodiment, microbes suitable for testing are bacteria from the family *Enterobacteriaceae*. In preferred embodiments, the compound is effective against a bacteria of a genus selected from the group consisting of: *Escherichia*, *Proteus*, *Salmonella*, *Klebsiella*, *Providencia*, *Enterobacter*, *Burkholderia*, *Pseudomonas*, *Aeromonas*, *Haemophilus*, *Yersinia*, *Neisseria*, and *Mycobacteria*.

In yet other embodiments, the microbes to be tested are Gram positive bacteria and are from a genus selected from the group consisting of: *Lactobacillus*, *Azorhizobium*, *Streptomyces*, *Pediococcus*, *Photobacterium*, *Bacillus*, *Enterococcus*, *Staphylococcus*, *Clostridium*, and *Streptococcus*.

In other embodiments, the microbes to be tested are fungi. In a preferred embodiment the fungus is from the genus *Mucor* or *Candida*, *e.g.*, *Mucor racmeosus* or *Candida albicans*.

In yet other embodiments, the microbes to be tested are protozoa. In a preferred embodiment the microbe is a malaria or cryptosporidium parasite.

VI. Transcription factor Modulating Compounds and Test Compounds

Compounds for testing in the instant methods can be derived from a variety of different sources and can be known or can be novel. In one embodiment, libraries of compounds are tested in the instant methods to identify transcriptional activation factor modulating compounds, e.g., HTH protein modulating compounds, AraC family polypeptide modulating compounds, MarA family polypeptide modulating compounds, etc. In another embodiment, known compounds are tested in the instant

methods to identify transcription factor modulating compounds (such as, for example, HTH protein modulating compounds, AraC family polypeptide modulating compounds, MarA family polypeptide modulating compounds, etc.). In an embodiment, compounds among the list of compounds generally regarded as safe (GRAS) by the Environmental Protection Agency are tested in the instant methods. In another embodiment, the transcription factors which are modulated by the modulating compounds are of prokaryotic microbes.

A recent trend in medicinal chemistry includes the production of mixtures of compounds, referred to as libraries. While the use of libraries of peptides is well established in the art, new techniques have been developed which have allowed the production of mixtures of other compounds, such as benzodiazepines (Bunin et al. 1992. *J. Am. Chem. Soc.* 114:10987; DeWitt et al. 1993. Proc. Natl. Acad. Sci. USA 90:6909) peptoids (Zuckermann. 1994. *J. Med. Chem.* 37:2678) oligocarbamates (Cho et al. 1993. Science. 261:1303), and hydantoins (DeWitt et al. supra). Rebek et al. have described an approach for the synthesis of molecular libraries of small organic molecules with a diversity of 104-105 (Carell et al. 1994. *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al. *Angew. Chem. Int. Ed. Engl.* 1994. 33:2061).

The compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the 'one-bead one-compound' library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. *Anticancer Drug Des.* 1997. 12:145).

Exemplary compounds which can be screened for activity include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, and natural product extract libraries. In one embodiment, the test compound is a peptide or peptidomimetic. In another, preferred embodiment, the compounds are small, organic non-peptidic compounds.

Other exemplary methods for the synthesis of molecular libraries can be found in the art, for example in: Erb et al. 1994. *Proc. Natl. Acad. Sci. USA* 91:11422; Horwell et al. 1996 *Immunopharmacology* 33:68; and in Gallop et al. 1994. *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner

USP '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*). Other types of peptide libraries may also be expressed, see, for example, U.S. Patents 5,270,181 and 5,292,646). In still another embodiment, combinatorial polypeptides can be produced from a cDNA library.

In other embodiments, the compounds can be nucleic acid molecules. In preferred embodiments, nucleic acid molecules for testing are small oligonucleotides. Such oligonucleotides can be randomly generated libraries of oligonucleotides or can be specifically designed to reduce the activity of a transcription factor, e.g., a HTH protein, a MarA family polypeptide, or an AraC family polypeptide. For example, in one embodiment, these oligonucleotides are sense or antisense oligonucleotides. In an embodiments, oligonucleotides for testing are sense to the binding site of a particular transcription factor, e.g., a MarA family polypeptide helix-turn-helix domain. Methods of designing such oligonucleotides given the sequences of a particular transcription factor polypeptide, such as a MarA family polypeptide, is within the skill of the art.

In yet another embodiment, computer programs can be used to identify individual compounds or classes of compounds with an increased likelihood of modulating a transcription factor activity, e.g., an HTH protein, a AraC family polypeptide, or a MarA family polypeptide activity. Such programs can screen for compounds with the proper molecular and chemical complementarities with a chosen transcription factor. In this manner, the efficiency of screening for transcription factor modulating compounds in the assays described above can be enhanced.

VII. Computer Modeling Techniques for Identifying Transcription factor Modulating Compounds

The invention also pertains to the use of molecular design techniques to design transcription factor modulating compounds, e.g., HTH protein modulating compounds, AraC family modulating compounds, MarA family modulating compounds, or MarA modulating compounds, which are capable of binding or interacting with one or more transcription factors (e.g., of a prokaryotic or eukaryotic organism). The invention pertains to both the transcription factor modulating compounds identified by the methods as well as the modeling methods, and compositions comprising the compounds identified by the methods.

In an embodiment, the invention pertains to a method of identifying transcription factor modulating compounds. The method includes obtaining the structure of a transcription factor of interest, and using GLIDE to identify a scaffold

which has an interaction energy score of –20 or less (e.g., -40 or less, e.g., -60 or less) with a portion of the transcription factor.

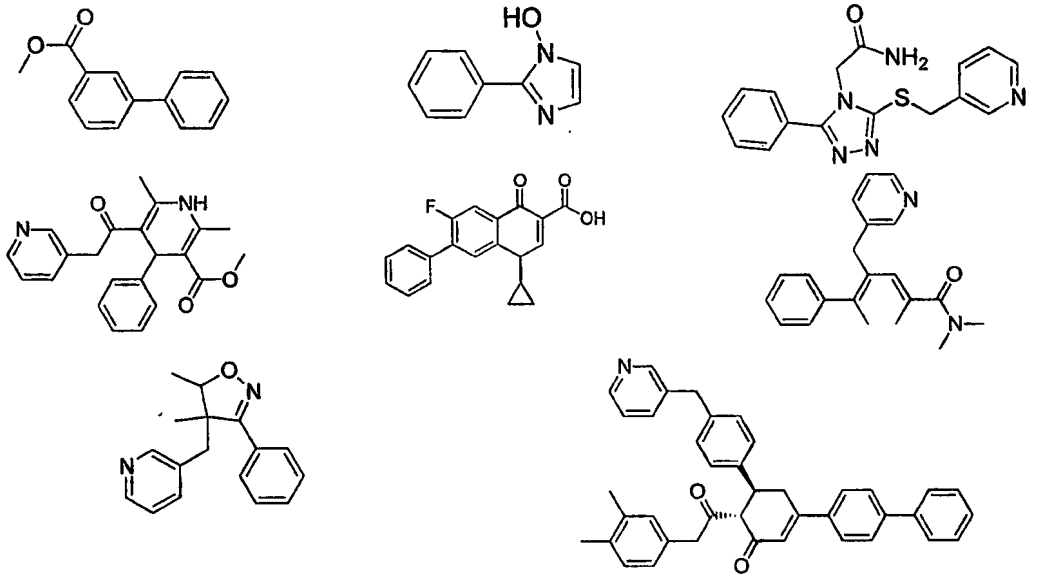
The invention pertains, at least in part, to a computational screening of small molecule databases for chemical entities or compounds that can bind in whole, or in part, to a transcription factor, such as a HTH protein, an AraC family polypeptide, a MarA family polypeptide, e.g., MarA. In this screening, the quality of fit of such entities or compounds to the binding site may be judged either by shape complementarity or by estimated interaction energy (Meng, E. C. *et al.*, 1992, *J. Coma. Chem.*, 13:505-524). Such a procedure allows for the screening of a very large library of potential transcription factor modulating compounds for the proper molecular and chemical complementarities with a selected protein or class or proteins. Transcription factor modulating compounds identified through computational screening can later be passed through the *in vivo* assays described herein as further screens. For example, a MarA inhibiting compound identified through computational screening could be tested for its ability to promote cell survival in a cell system containing a counterselectable marker under the control a MarA activated promoter. The promotion of cell survival in the foregoing assay would be indicative of a compound that inhibits MarA's activity as a transcriptional activator. Other suitable assays are described in the Examples and through the specification.

The crystal structures of both MarA (PDB ID code 1BL0) and its homolog Rob (PDB ID code 1DY5) are available in the Protein Data Bank (http://www.rcsb.org/pdb/). These structures were used to identify sites on the proteins that could be targeted by small molecule chemical inhibiting compounds. A total of at least eight potential small molecule binding sites on MarA (Table 2) and four sites on Rob (Table 3) were identified as potential small molecule binding sites. The invention pertains, at least in part, to MarA modulating compounds which interact with any one of the following sites of MarA (based on the sequence given in SEQ ID NO. 2).

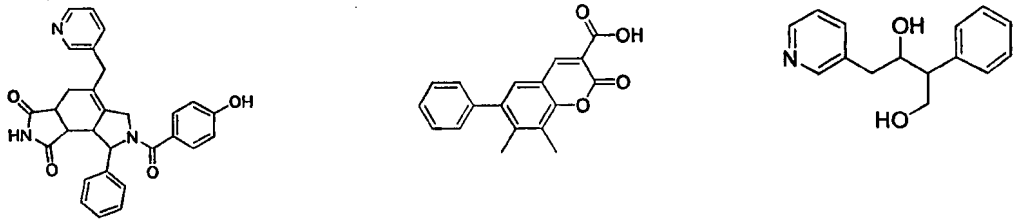
Table 2

Site Number	Residues (based on full length MarA)	Site Label
1	42 to 50	R46 Major Groove
2	54 to 62	L56 HTH core
3	55 to 65	R61 Minor Groove
4	15 to 25	W19
5	14 to 25	E21
6	24 to 35	L28
7	76 to 83	P78
8	106 to 112	R110

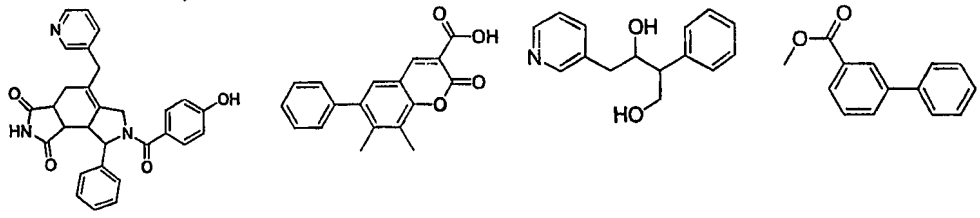
The GLIDE docking method was then used to fit combinatorial chemistry scaffolds into these sites and an interaction energy was calculated for each. Eight scaffolds were predicted to bind to site 1, encompassing amino acids tryptophan 42 to lysine 50, with an interaction energy score of –60 or less. These scaffolds are shown below:



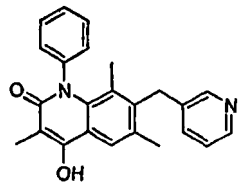
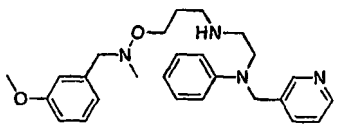
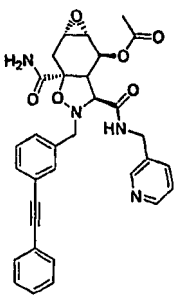
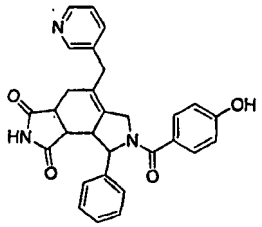
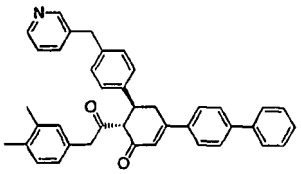
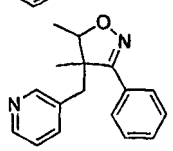
Three scaffolds were identified for site 2 of MarA (e.g., residues histidine 54 to serine 62).



Four scaffolds were identified for MarA site 3, (e.g., residues serine 55 to methionine 65):



Six scaffolds were identified for site 6 (e.g., residues leucine 24 to glutamate 35).

These scaffolds were then used to search the CambridgeSoft ACX-SC database of over 600,000 non-proprietary chemical structures and the number of chemicals similar to the scaffolds was determined.

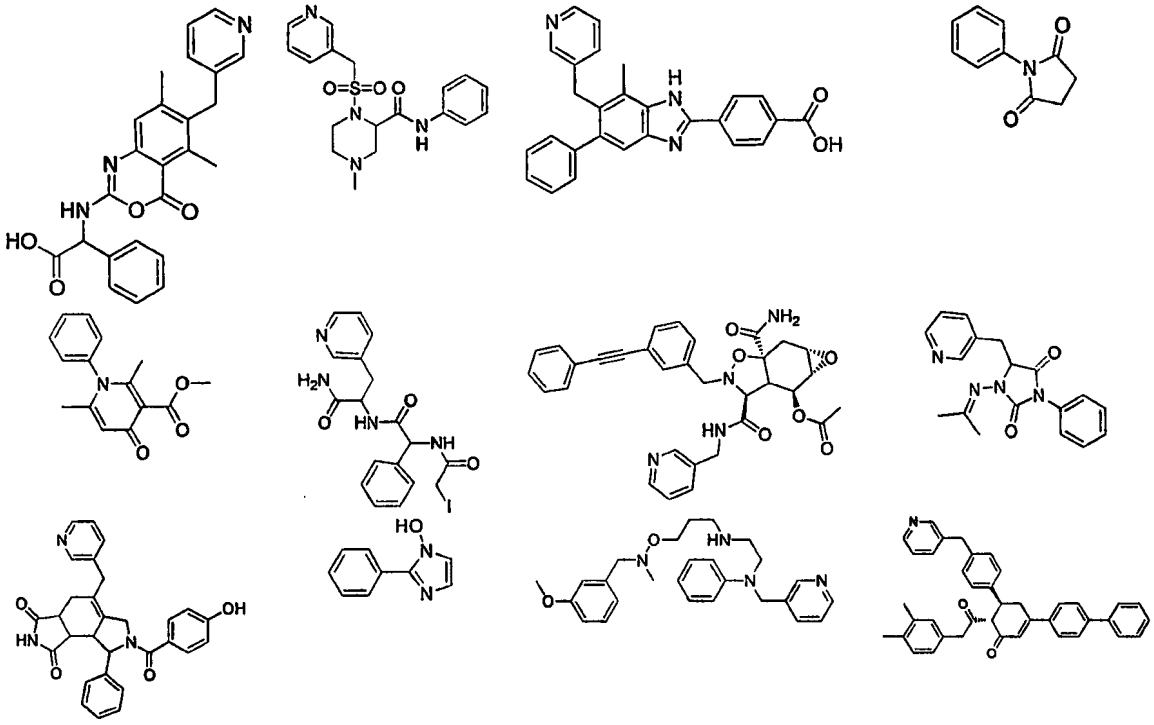
The term "scaffold" includes the compounds identified by the computer modeling program. These compounds may or may not be themselves transcription factor modulating compounds. An ordinarily skilled artisan will be able to analyze a scaffold obtained from the computer modeling program and modify the scaffold such that the resulting compounds have enhanced chemical properties over the initial scaffold compound, e.g., are more stable for administration, less toxic, have enhanced affinity for a particular transcription factor, etc. The invention pertains not only to the scaffolds identified, but also the transcription factor modulating compounds which are developed using the scaffolds.

Table 3 lists portions of Rob which were identified as possible interaction sites for a modulating compound. The invention pertains, at least in part, to any compounds modeled to bind to these regions of Rob. The numbering corresponds to that given in SEQ ID NO. 4.

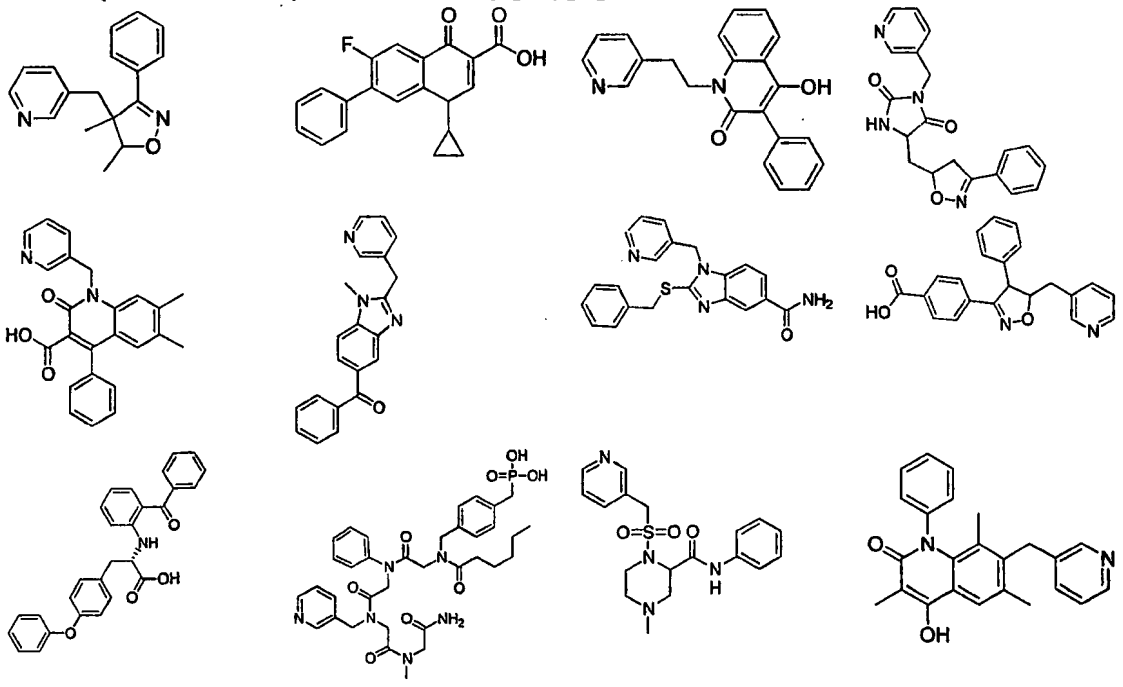
Table 3

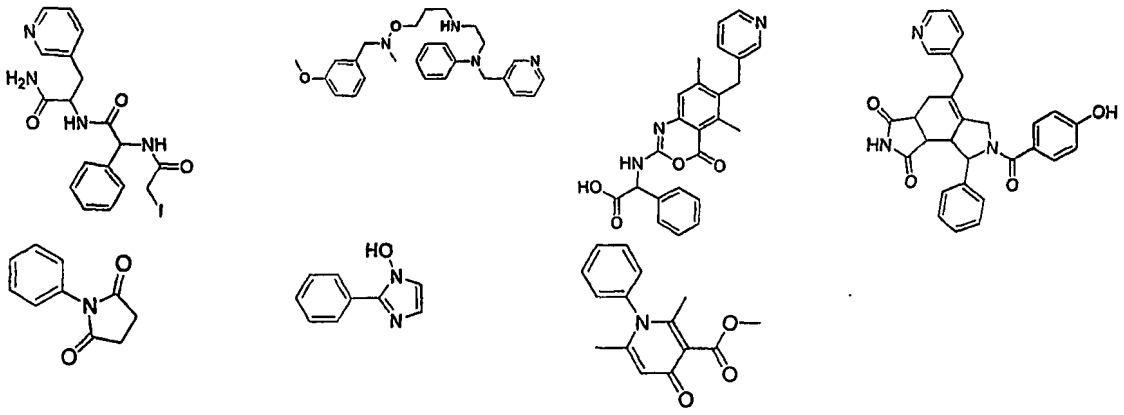
Site Number	Residues (based on full length Rob)	Site Label
1	37 to 45	R40 Major Groove
2	43 to 54	I50 HTH Core
3	51 to 60	R55 Minor Groove
4	10 to 20	W13

These scaffolds were identified as possible modulating compounds which with site 1 of Rob (residues 37-45), a MarA family polypeptide.



These scaffolds were identified as small molecules that may interact with site 2 of Rob (residues 43-52), a MarA family polypeptide.





The design of compounds that bind to, modulate, or inhibit transcription factors, generally involves consideration of two factors. First, the compound must be capable of physically and structurally associating with a particular transcription factor. Non-covalent molecular interactions important in the association of a transcription factor with a modulating compound include hydrogen bonding, van der Waals and hydrophobic interactions.

Second, the modulating compound must be able to assume a conformation that allows it to associate with the selected transcription factor. Although certain portions of the inhibiting compound will not directly participate in this association with the transcription factor, those portions may still influence the overall conformation of the molecule. This, in turn, may have a significant impact on potency. Such conformational requirements include the overall three-dimensional structure and orientation of the chemical entity or compound in relation to all or a portion of the binding site, *e.g.*, active site or accessory binding site of a particular transcription factor such as MarA, or the spacing between functional groups of a compound comprising several chemical entities that directly interact with the particular transcription factor.

In a further embodiment, the potential modulating effect of a chemical compound on a selected transcription factor (e.g., a HTH protein, a AraC family polypeptide, a MarA family polypeptide, e.g., MarA) is analyzed prior to its actual synthesis and testing by the use of computer modeling techniques. If the theoretical structure of the given compound suggests insufficient interaction and association between it and the selected transcription factor, synthesis and testing of the compound is avoided. However, if computer modeling indicates a strong interaction, the molecule may then be synthesized and tested for its ability to bind to the selected transcription factor and modulate the transcription factor's activity.

A transcription factor modulating compound or other binding compound (e.g., an HTH protein modulating compound, an AraC family polypeptide modulating compound, or a MarA family inhibiting compound, e.g., a MarA inhibiting compound)

may be computationally evaluated and designed by screening and selecting chemical entities or fragments for their ability to associate with the individual small molecule binding sites or other areas of a transcription factor.

One skilled in the art may use one of several methods to screen chemical entities or fragments for their ability to associate with a selected transcription factor and more particularly with the individual small molecule binding sites of the particular transcription activation factor. This process may begin by visually inspecting the structure of the transcription factor on a computer screen based on the atomic coordinates of the transcription factor crystals. Selected chemical entities may then be positioned in a variety of orientations, or docked, within an individual binding site of the transcription factor. Docking may be performed using software such as Quanta and Sybyl, followed by energy minimization with standard molecular mechanics forcefields or dynamics with programs such as CHARMM (Brooks, B. R. et al., 1983, *J. Comp. Chem.*, 4:187-217) or AMBER (Weiner, S. J. et al., 1984, *J. Am. Chem. Soc.*, 106:765-784).

Specialized computer programs may also assist in the process of selecting molecules that bind to a selected transcription factor, (e.g., an HTH protein, an AraC family polypeptide, or a MarA family polypeptide, e.g., MarA). The programs include, but are not limited to:

1. GRID (Goodford, P. J., 1985, "A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules" *J. Med. Chem.*, 28:849-857 GRID is available from Oxford University, Oxford, UK.

2. AUTODOCK (Goodsell, D. S. and A. J. Olsen, 1990, "Automated Docking of Substrates to Proteins by Simulated Annealing" Proteins: Structure. *Function, and Genetics*, 8:195-202. AUTODOCK is available from Scripps Research Institute, La Jolla, Calif. AUTODOCK helps in docking inhibiting compounds to a selected transcription factor in a flexible manner using a Monte Carlo simulated annealing approach. The procedure enables a search without bias introduced by the researcher.

3. MCSS (Miranker, A. and M. Karplus, 1991, "Functionality Maps of Binding Sites: A Multiple Copy Simultaneous Search Method." Proteins: Structure, Function and Genetics, 11:29-34). MCSS is available from Molecular Simulations, Burlington, Mass.

4. MACCS-3D (Martin, Y. C., 1992, J. Med. Chem., 35:2145-2154) is a 3D database system available from MDL Information Systems, San Leandro, Calif.

5. DOCK (Kuntz, I. D. *et al.*, 1982, "A Geometric Approach to Macromolecule-Ligand Interactions" *J. Mol. Biol.*, 161:269-288). DOCK is available from University of California, San Francisco, Calif. DOCK is based on a description of the negative image of a space-filling representation of the molecule (*i.e.* the selected transcription factor) that should be filled by the inhibiting compound. DOCK includes a force-field for energy evaluation, limited conformational flexibility and consideration of hydrophobicity in the energy evaluation.

6. MCDLNG (Monte Carlo De Novo Ligand Generator) (D. K. Gehlhaar, *et al.* 1995. *J. Med. Chem.* 38:466-472). MCDLNG starts with a structure (i.e. an X-ray crystal structure) and fills the binding site with a close packed array of generic atoms. A Monte Carlo procedure is then used to randomly: rotate, move, change bond type, change atom type, make atoms appear, make bonds appear, make atoms disappear, make bonds disappear, etc. The energy function used by MCDLNG favors the formation of rings and certain bonding arrangements. Desolvation penalties are given for heteroatoms, but heteroatoms can benefit from hydrogen bonding with the binding site.

In an embodiment of the invention, docking is performed by using the Affinity program within InsightII (Molecular Simulations Inc., 1996, San Diego, Calif., now Accelrys Inc.). Affinity is a suite of programs for automatically docking a ligand (i.e. a transcription factor modulating compound) to a receptor (i.e. a transcription factor). Commands in Affinity automatically find the best binding structures of the ligand to the receptor based on the energy of the ligand/receptor complex. As described below, Affinity allows for the simulation of flexible-flexible docking.

Affinity consists of two commands, **GridDocking** and **fixedDocking**, under the new pulldown **Affinity** in the **Docking** module of the Insight II program. Both commands use the same, Monte Carlo type procedure to dock a guest molecule (*i.e.* HTH protein modulating compound) to a host (i.e., a transcription factor). They also share the feature that the "bulk" of the receptor (i.e. transcription factor), defined as atoms not in the binding (active) site specified, is held rigid during the docking process,

while the binding site atoms and ligand atoms are movable. The commands differ, however, in their treatment of nonbond interactions. In **GridDocking**, interactions between bulk and movable atoms are approximated by the very accurate and efficient molecular mechanical/grid (MM/Grid) method developed by Luty *et al.* 1995. *J. Comp. Chem.* 16:454, while interactions among movable atoms are treated exactly. **GridDocking** also includes the solvation method of Stouten *et al.* 1993. Molecular Simulation 10:97. On the other hand, the **fixedDocking** command computes nonbond interactions using methods in the Discover program (cutoff methods and the cell multipole method) and it does not include any solvation terms.

Affinity does not, generally, require any intervention from the user during the docking. It automatically moves the ligand (i.e. modulating compound), evaluates energies, and checks if the structure is acceptable. Moreover, the ligand and the binding site of the receptor (i.e. the selected transcription modulator) are flexible during the search.

Most of the docking methods in the literature are based on descriptors or empirical rules (for a review see Kuntz *et al.* 1994. *Acc. Chem. Res.* 27:117. These include DOCK (Kuntz *et al.* 1982. *J. Mol. Biol.* 161:269., Shoichet *et al.* 1992. *J. Compt. Chem.* 13:380., Oshiro *et al.* 1995. *J. Comp. Aided Molec. Design* 9:113.), CAVEAT (Bartlett *et al.* 1989. "Chemical and Biological Problems in Molecular Recognition" Royal Society of Chemistry: Cambridge, pp. 182-196., Lauri & Bartlett. 1994. *J. Comput. Aided Mol. Design* 8:51), FLOG (Miller *et al.* 1994. *J. Comp. Aided Molec. Design* 8:153), and PRO_LIGAND (Clark *et al.* 1995. *J. Comp. Aided Molec. Design* 9:13), to name a few. Affinity differs from these methods in several aspects.

First, it uses full molecular mechanics in searching for and evaluating docked structures. In contrast descriptor-based methods use empirical rules which usually take into account only hydrogen bonding, hydrophobic interactions, and steric effects. This simplified description of ligand/receptor interaction is insufficient in some cases. For example, Meng *et al.* 1992. *J. Compt. Chem.* 13:505 studied three scoring methods in evaluating docked structures generated by DOCK. They found that only the forcefield scores from molecular mechanics correctly identify structures closest to experimental binding geometry, while scoring functions that consider only steric factors or only electrostatic factors are less successful. Note that in the study by Meng *et al.* 1992. *J. Compt. Chem.* 13:505, docking was still performed using descriptors. Affinity, on the other hand, uses molecular mechanics in both docking and scoring and is therefore more consistent.

Second, in Affinity, while the bulk of the receptor is fixed, the defined binding site is free to move, thereby allowing the receptor to adjust to the binding of different ligands or different binding modes of the same ligand. By contrast, almost all of the descriptor-based methods fix the entire receptor.

Third, the ligand itself is flexible in Affinity which permits different conformations of a ligand (*i.e.* transcription factor modulating compound) to be docked to a receptor (i.e. transcription factor). Recently Oshiro *et al.* (1995 *J. Comp. Aided Molec. Design* 9;113) extended DOCK to handle flexible ligands. FLOG is also able to treat flexible ligand by including different conformations for each structure in the database (Miller et al. 1995. *J. Comp. Aided Molec. Design.* 8:153). Most other methods are limited to rigid ligands.

There are also a few energy based docking methods (Kuntz *et al.* 1994. *Acc. Chem Res.* 27:117). These methods use either molecular dynamics (notably simulated annealing) or Monte Carlo methods. For example, Caflisch *et al.* 1992. *Proteins: Struct. Funct. and Genetics* 13:223) developed a two step procedure for docking flexible ligands. In their procedure, ligand is first docked using a special energy function designed to remove bad contact between the ligand and the receptor efficiently. Then Monte Carlo minimization (Li & Scheraga. 1987. *Proc. Natl. Acad. Sci. U.S.A.* 84:6611) is carried out to refine the docked structures using molecular mechanics. Hart and Read. 1992. *Proteins: Struct. Funct. and Genetics* 13:206 also employ two steps to dock ligands. They use a score function based on receptor geometry to approximately dock ligands in the first step, and then use Monte Carlo minimization similar to that of Caflisch et al. 1992. *Proteins: Struct. Funct. and Genetics* 13:223 for the second step. The method by Mizutani *et al.* (1994. *J. Mol. Biol.* 243:310) is another variation of this two step method.

Affinity uses a Monte Carlo procedure in docking ligands, but there are important distinctions over the prior art methods. First, the Monte Carlo procedure in Affinity can be used in conjunction either with energy minimization (to mimic the Monte Carlo minimization method of Li & Scheraga. 1987. *Proc. Natl. Acad. Sci. U.S.A.* 84:6611) or with molecular dynamics (to mimic the hybrid Monte Carlo method, Clamp et al. 1994. *J. Comput. Chem.* 15:838, or the smart Monte Carlo method, Senderowitz et al. 1995. *J. Am. Chem. Soc.* 117:8211). This flexibility allows Affinity to be applied to a variety of docking problems. Second, in the initial screening of docked structures, Affinity employs energy differences obtained from molecular mechanics, while the methods discussed above use empirical rules or descriptors. Therefore, Affinity is more consistent in that it uses molecular mechanics in both initial screening and final refinement of docked structures. Third, Affinity allows the binding

site of the receptor to relax, while the methods discussed above fix the entire receptor. Fourth, Affinity employs two new nonbond techniques which are both accurate and efficient to make docking practical. One is the Grid/MM method of Luty *et al.* which represents the bulk of the receptor by grids (Luty *et al.* 1995. *J. Comp. Chem.* 16:454). This method is 10-20 times faster than the no-cutoff method with almost no loss in accuracy. It also incorporates the solvation method of Stouten *et al.* (1993. *Molecular Simulation* 10:97). The other is the cell multipole method. This method is about 50% slower than the Grid/MM method, but it does not require grid setup. Thus, a typical docking calculation takes about 1-3 hours of CPU time on an Indigo R4400 workstation.

Once suitable chemical fragments have been selected, they can be assembled into a single compound or inhibiting compound. Assembly may be proceed by visual inspection of the relationship of the fragments to each other on a three-dimensional image display on a computer screen in relation to the structure coordinates of a particular transcription factor, e.g., MarA. This may be followed by manual model building using software such as Quanta or Sybyl.

Useful programs to aid one of skill in the art in connecting the individual chemical fragments include:

1. 3D Database systems such as MACCS-3D (MDL Information Systems, San Leandro, Calif. This area is reviewed in Martin, Y. C., 1992, "3D Database Searching in Drug Design", *J. Med. Chem.*, 35, pp. 2145-2154).

2. CAVEAT (Bartlett, P. A. et al, 1989, "CAVEAT: A Program to Facilitate the Structure-Derived Design of Biologically Active Molecules". In Molecular Recognition in Chemical and Biological Problems", Special Pub., Royal Chem. Soc., 78, pp. 182-196). CAVEAT is available from the University of California, Berkeley, Calif. CAVEAT suggests inhibiting compounds to MarA based on desired bond vectors.

3. HOOK (available from Molecular Simulations, Burlington, Mass.). HOOK proposes docking sites by using multiple copies of functional groups in simultaneous searches.

In another embodiment, transcription factor modulating compounds may be designed as a whole or "de novo" using either an empty active site or optionally including some portion(s) of a known inhibiting compound(s). These methods include:

1. LUDI (Bohm, H.-J., "The Computer Program LUDI: A New Method for the De Novo Design of Enzyme Inhibiting compounds", J. ComR. Aid. Molec. Design, 6, pp. 61-78 (1992)). LUDI is available from Biosym Technologies, San Diego, Calif. LUDI is a program based on fragments rather than on descriptors. LUDI proposes somewhat larger fragments to match with the interaction sites of a macromolecule and scores its hits based on geometric criteria taken from the Cambridge Structural Database (CSD), the Protein Data Bank (PDB) and on criteria based on binding data. LUDI is a library based method for docking fragments onto a binding site. Fragments are aligned with 4 directional interaction sites (lipophilic-aliphatic, lipophilic-aromatic, hydrogen donor, and hydrogen acceptor) and scored for their degree of overlap. Fragments are then connected (i.e. a linker of the proper length is attached to each terminal atom in the fragments). Note that conformational flexibility can be accounted for only by including multiple conformations of a particular fragment in the library.

2. LEGEND (Nishibata, Y. and A. Itai, Tetrahedron, 47, p. 8985 (1991)). LEGEND is available from Molecular Simulations, Burlington, Mass.

3. CoMFA (Conformational Molecular Field Analysis) (J. J. Kaminski. 1994. *Adv. Drug Delivery Reviews* 14:331-337.) CoMFA defines 3-dimensional molecular shape descriptors to represent properties such as hydrophobic regions, sterics, and electrostatics. Compounds from a database are then overlaid on the 3D pharmacophore model and rated for their degree of overlap. Small molecule databased that be searched include: ACD (over 1,000,000 compounds), Maybridge (about 500,000 compounds), NCI (about 500,000 compounds), and CCSD. In measuring the goodness of the fit, molecules can either be fit to the 3D molecular shape descriptors or to the active conformation of a known inhibiting compound.

4. LeapFrog (available from Tripos Associates, St. Louis, Mo.).

FlexX (© 1993-2002 GMD German National Research Center for Information Technology; Rarey, M. *et al J. Mol. Biol.*, 261:407-489) is a fast, flexible docking method that uses an incremental construction algorithm to place ligands into

and active site of the transcription factor. Ligands (e.g., transcription factor modulating compounds) that are capable of "fitting" into the active site are then scored according to any number of available scoring schemes to determine the quality of the complimentarity between the active site and ligand.

Other molecular modeling techniques may also be employed in accordance with this invention. See, e.g., Cohen, N. C. et al., "Molecular Modeling Software and Methods for Medicinal Chemistry, J. Med. Chem., 33, pp. 883-894 (1990). See also, Navia, M. A. and M. A. Murcko, "The Use of Structural Information in Drug Design", Current Opinions in Structural Biology, 2, pp. 202-210 (1992).

Candidate transcription factor modulating compounds can be evaluated for their modulating, e.g., inhibitory, activity using conventional techniques which may involve determining the location and binding proximity of a given moiety, the occupied space of a bound inhibiting compound, the deformation energy of binding of a given compound and electrostatic interaction energies. Examples of conventional techniques useful in the above evaluations include, but are not limited to, quantum mechanics, molecular dynamics, Monte Carlo sampling, systematic searches and distance geometry methods (Marshall, G. R., 1987, *Ann. Ref. Pharmacol. Toxicol.*, 27:193). Examples of computer programs for such uses include, but are not limited to, Gaussian 92, revision E2 (Gaussian, Inc. Pittsburgh, Pennsylvania), AMBER version 4.0 (University of California, San Francisco), QUANTA/CHARMM (Molecular Simulations, Inc., Burlington, Mass.), and Insight II/Discover (Biosym Technologies Inc., San Diego, Calif.). These programs may be implemented, for example, using a Silicon Graphics Indigo2 workstation or IBM RISC/6000 workstation model 550. Other hardware systems and software packages will be known and of evident applicability to those skilled in the art.

Once a compound has been designed and selected by the above methods, the efficiency with which that compound may bind to a particular transcription factor may be tested and optimized by computational evaluation. An effective transcription factor modulating compound should demonstrate a relatively small difference in energy between its bound and free states (i.e., a small deformation energy of binding). Transcription factor modulating compounds may interact with the selected transcription factor in more than one conformation that is similar in overall binding energy. In those cases, the deformation energy of binding may be taken to be the difference between the energy of the free compound and the average energy of the conformations observed when the inhibiting compound binds to the enzyme.

A compound designed or selected as interacting with a selected transcription factor, e.g., a MarA family polypeptide, e.g., MarA, may be further computationally optimized so that in its bound state it would preferably lack repulsive electrostatic interaction with the target enzyme. Such non-complementary (e.g., electrostatic) interactions include repulsive charge-charge, dipole-dipole and charge-dipole interactions. Specifically, the sum of all electrostatic interactions between the modulating compound and the enzyme when the modulating compound is bound to the selected transcription factor, preferably make a neutral or favorable contribution to the enthalpy of binding.

Specific computer software is available in the art to evaluate compound deformation energy and electrostatic interaction. Examples of programs designed for such uses include: Gaussian 92, revision C [M. J. Frisch, Gaussian, Inc., Pittsburgh, Pa. ©1992]; AMBER, version 4.0 [P. A. Kollman, University of California at San Francisco, ©1994]; QUANTA/CHARMM [Molecular Simulations, Inc., Burlington, Mass. ©1994]; and Insight II/Discover (Biosysm Technologies Inc., San Diego, Calif. ©1994). These programs may be implemented, for instance, using a Silicon Graphics workstation, IRIS 4D/35 or IBM RISC/6000 workstation model 550. Other hardware systems and software packages will be known to those skilled in the art.

Once a transcription factor modulating compound has been optimally selected or designed, as described above, substitutions may then be made in some of its atoms or side groups in order to improve or modify its binding properties. Initial substitutions are preferable conservative, i.e., the replacement group will have approximately the same size, shape, hydrophobicity and charge as the original group. Substitutions known in the art to alter conformation should be avoided. Such substituted chemical compounds may then be analyzed for efficiency of fit to the selected transcription factor by the same computer methods described above.

Computer programs can be used to identify unoccupied (aqueous) space between the van der Waals surface of a compound and the surface defined by residues in the binding site. These gaps in atom-atom contact represent volume that could be occupied by new functional groups on a modified version of the lead compound. More efficient use of the unoccupied space in the binding site could lead to a stronger binding compound if the overall energy of such a change is favorable. A region of the binding pocket which has unoccupied volume large enough to accommodate the volume of a group equal to or larger than a covalently bonded carbon atom can be identified as a promising position for functional group substitution. Functional group substitution at this region can constitute substituting something other than a carbon atom, such as oxygen. If the volume is large enough to accommodate a group larger than a carbon

atom, a different functional group which would have a high likelihood of interacting with protein residues in this region may be chosen. Features which contribute to interaction with protein residues and identification of promising substitutions include hydrophobicity, size, rigidity and polarity. The combination of docking, K_i estimation, and visual representation of sterically allowed room for improvement permits prediction of potent derivatives.

Similarity Screening

Once a transcription factor modulating compound has been selected or designed, computational methods to assess its overall likeness or similarity to other molecules can be used to search for additional compounds with similar biochemical behavior. In such a way, for instance, HTS derived hits can be tested to assure that they are bona fide ligands against a particular active site, and to eliminate the possibility that a particular hit is an artifact of the screening process. There are currently several methods and approaches to determine a particular compound's similarity to members of a virtual database of compounds. One example is the OPTISIM methodology that is distributed in the Tripos package, SYBYL (© 1991-2002 Tripos, Inc., St. Louis, MO). OPTISIM exploits the fact that each 3-dimensional representation of a molecule can be broken down into a set of 2-dimensional fragments and encoded into a pre-defined binary string. The result is that each compound within a particular set is represented by a unique numerical code or fingerprint that is amenable to mathematical manipulations such as sorting and comparison. OPTISIM is automated to calculate and report the percent difference in the fingerprints of the respective compounds for instance according to the using a formalism known as the Tanimoto coefficient. For instance, a compound that is similar in structure to another will share a high coefficient. Large virtual databases of commercially available compounds or of hypothetical compounds can be quickly screened to identify compounds with high Tanimoto coefficient.

CoMFA/QSAR

Once a series of similar transcription factor modulating compounds has been identified and expanded by the methods described, their experimentally determined biological activities can be correlated with their structural features using a number of available statistical packages. In a typical project within the industry, the CoMFA (COmparative Molecular Field Analysis) and QSAR (Quantitative Structure Activity Relationship) packages within the SYBYL suite of programs (Tripos Associates, St. Louis, MO) are utilized. In CoMFA, a particular series of compounds with measured activities are co-aligned in a manner that is believed to emulate their arrangement as

they interact with the active site. A 3-dimensional lattice, or grid is then constructed to encompass the collection of the so-aligned compounds. At each point on the lattice, an evaluation of the potential energy is determined and tabulated-typically potentials that simulate the electronic and steric fields are determined, but other potential functions are available. Using the statistical methods such as PLS (Partial Least Squares), correlation between the measured activities and the potential energy values at the grid-points can be determined and summed in a linear equation to produce the overall molecular correlation or QSAR model. A particularly useful feature in CoMFA is that the individual contribution for each grid-point is known; the importance of the grid points upon the overall correlation can be visualized graphically in what is referred to as a CoMFA field. When this field is combined with the original compound alignment, it becomes a powerful tool to rationalize the activities of the individual compounds from whence the model was derived, and to predict how chemical modification of a reference compound would be effected. As an example, a QSAR model was developed for a set of 92 benzodiazepines using the method described above. A representative CoMFA field is shown in Figure 4; the region delineated by wire mesh (adjacent to the referenced triazinoxazepine) is the region where chemical modification characterized by increasing steric bulk would lead to favorable effects in transcription factor modulation.

The invention pertains, per se, to not only the methods for identifying the transcription factor modulating compounds, but to the compounds identified by the methods of the invention as well as methods for using the identified compounds.

VIII. MarA family Modulating Compounds, and Methods of Use thereof

In an embodiment, the invention pertains to methods for modulating a transcription factor, e.g., an HTH protein, an AraC family polypeptide, or a MarA family polypeptide. The method includes contacting the transcription factor, e.g., a MarA family polypeptide, with a transcription factor modulutating compound of the formula (I):

A-E (I)

wherein A is a polar moiety, E is a hydrophobic moiety, and pharmaceutically acceptable salts thereof. The transcription factor modulating compound, e.g., a MarA family modulating compound, may comprise one or more polar moieties and/or one or more hydrophobic moieties.

In another embodiment, the invention pertains to methods for reducing antibiotic resistance of a microbial cell. The method includes contacting the cell with a transcription factor modulating compound, e.g., a MarA family modulating compound, such that the antibiotic resistance of the cell is reduced.

In another embodiment, the invention pertains to inhibiting transcription, comprising contacting a transcription factor with a transcription factor modulating compound, such that transcription is inhibited. In a further embodiment, the transcription of a prokaryotic cell is inhibited. In another further embodiment, the transcription factor modulating compound is a compound of anyone of formulae (I)-(X).

The term "antibiotic resistance" includes resistance of a microbial cell to a antibiotic compound, especially an antibiotic compound which had been previously used to treat similar microbial organisms successfully.

The term "polar moiety" includes moieties with at least one heterocycle. It also includes moieties such as, but not limited to, hydroxyl, halogens, thioethers, carboxylic acids, metals (*e.g.* alkali, alkaline, Au, Hg, Ag, Mn, Co, Cu, Zn, etc.), nitro, amino, alkoxy, and other moieties which allow the compound to perform its intended function. The term "polar moiety" includes moieties which allow the transcription factor modulating compound to perform its intended function, *e.g.*, modulate a transcription factor, e.g., an AraC family polypeptide or a MarA family polypeptide. A heterocyclic polar moiety may comprise one or more rings, one or more of which may be aromatic. In an embodiment, one or more rings of the polar moiety are fused. The heterocyclic polar moiety may also be bicyclic.

The heterocyclic polar moiety may comprise one or more nitrogen, sulfur, or oxygen atoms. Examples of heterocycles include benzodioxazole, benzofuran, benzoimidazole, benzoxazole, benzothiazole, benzothiophene, chromenone, deazapurine, furan, imidazole, imidazopyridine, indole, indolizine, isooxazole, isothiaozole, isoquinoline, methylenedioxyphenyl, napthridine, oxazole, purine, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrollidine, quinoline, tetrazole, thiazole, thiophene, triazole, and triazoletetrazole.

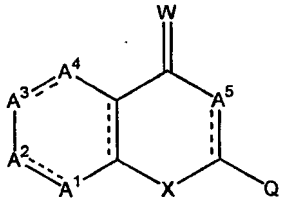
Furthermore, the polar moiety may be substituted when chemically feasible. For example, the polar moiety may be substituted with one or more substituents such as alkyl, alkenyl, alkynyl, halogen, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, arylcarbonyl, alkoxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino,

arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moiety. Examples of substituents also include nitro, alkoxy, aryl, amidyl, ester, thioester, alkyl (*e.g.*, methyl, ethyl, propyl, butyl, pentyl, etc.), araalkyl (*e.g.*, substituted or unsubstituted benzyl), hydroxy, halogen (*e.g.*, fluorine, bromine, chlorine, iodine, etc.).

The term "hydrophobic moiety" includes moieties such as which allow the transcription factor modulating compound (e.g., an HTH protein modulating compound, an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) to perform its intended function, *e.g.*, modulate a transcription factor. Examples of hydrophobic moieties include, for example, hydrogen, alkyl, alkenyl, alkynyl, and aryl moieties. The hydrophobic moieties may be unsubstituted or substituted, if chemically feasible (*e.g.*, not hydrogen). In an embodiment, the hydrophobic moiety is substituted or unsubstituted phenyl. Examples of substituents include alkyl, alkenyl, alkynyl, alkoxy, halogen, amino, thiol, hydroxy, nitro, aryl, and heteroaryl. The substituents can be substituted or unsubstituted. In an embodiment, the phenyl hydrophobic moiety is *para*-substituted, *e.g.*, alkyl (methyl, ethyl, propyl, butyl, pentyl, etc.), halogen (*e.g.*, fluorine, bromine, chlorine, iodine, etc.), hydroxy, substituted.

In another embodiment, the hydrophobic moiety is heterocyclic. Examples of heterocyclic hydrophobic moieties include imidazopyridine, quinolinyl, pyridinyl, *etc.*

In one embodiment, the transcription factor modulating compound (e.g., MarA family polypeptide modulating compound, AraC family polypeptide modulating compound, etc.) is of the formula (VII):



(VII)

wherein

W is NH, O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^1, O, or S;

A^2 is C-Z^2, O, or S;

A^3 is C-Z^3, O, or S;

A^4 is C-Z^4, O, or S;

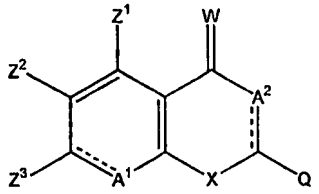
A^5 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, and Z^4 are each independently selected from the group consisting of hydrogen, alkoxy, hydroxy, halogen, and alkyl;

Z^5 is hydrogen, alkoxy, hydroxy, halogen, alkyl, or carbonyl;

Q is hydrogen, alkyl, alkenyl, alkynyl, halogen, hydroxy, aryl, and pharmaceutically acceptable salts thereof.

In yet another embodiment, the transcription factor modulating compound (e.g., the MarA family polypeptide modulating compound, AraC family polypeptide modulating compound, etc.) is of the formula (II):



(II)

wherein

W is O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^4, O, or S;

A^2 is C-Z^5, or N-Z^5;

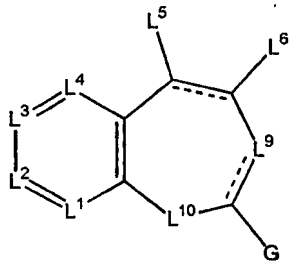
Z^1, Z^2, Z^3, Z^4 and Z^5 are each independently hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, or cyano;

Z^3 is hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, nitro, cyano, carbonyl, or thiocarbonyl;

Q is an aromatic or heterocyclic moiety, and pharmaceutically acceptable salts thereof.

In a further embodiment, W may be oxygen and X may be oxygen. Furthermore, A^1 and A^2may be C-Z^4 and C-Z^5, respectively. Examples of Z^4 and Z^5include hydrogen and hydroxy. Examples of Z^1 and Z^2 include hydrogen and hydroxy. Other examples of Z^2 also include halogen, *e.g.*, fluorine, chlorine, bromine, and iodine. Examples of Z^3 include, for example, hydrogen, alkoxy and hydroxy. Examples of Q include substituted and unsubstituted phenyl. The phenyl may be *para*-substituted. Examples of substituents include hydroxyl, halogen (*e.g.*, fluorine, bromine, chlorine, iodine, etc.), amino, alkyl (*e.g.*, methyl, ethyl, propyl, butyl, pentyl, etc.), nitro, cyano, etc. In an embodiment, the transcription factor modulating compound is a MarA modulating compound, and in a further embodiment, a MarA inhibiting compound.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (VIII):



(VIII)

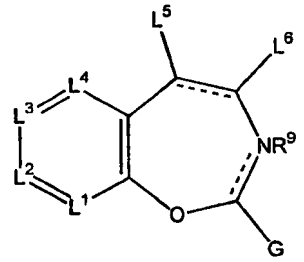
wherein:

G is a substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen;

L^1, L^2, L^3, L^4, L^9 and L^{10} are each independently oxygen, sulfur, substituted or unsubstituted nitrogen, and substituted or unsubstituted carbon; and

L^5 and L^6 are each independently hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, acyl, heterocyclic, amino, nitro, hydroxy, cyano, alkoxy, or aryl, and L^5 and L^6 may optionally be linked with a chain of one to six atoms to form a fused ring, and pharmaceutically acceptable salts thereof.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (IX):



(IX)

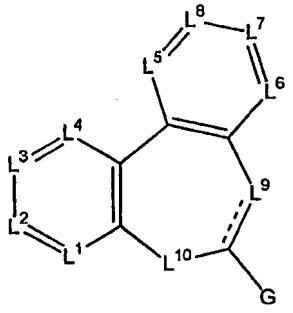
wherein:

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen;

L^1, L^2, L^3, and L^4 are each independently oxygen, sulfur, substituted or unsubstituted nitrogen, and substituted or unsubstituted carbon; and

R^9, L^5 and L^6 are each independently hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, acyl, heterocyclic, amino, nitro, hydroxy, cyano, alkoxy, or aryl, and L_5 and L_6 may optionally be linked with a chain of one to six atoms to form a fused ring, and pharmaceutically acceptable salts thereof.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (III):



(III)

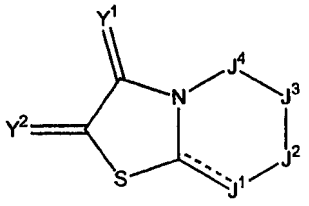
wherein

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen; and

L^1, L^2, L^3, L^4, L^5, L^6, L^7, L^8, L^9, and L^{10} are each independently oxygen, substituted or unsubstituted nitrogen, sulfur and or substituted or unsubstituted carbon, and pharmaceutically acceptable salts thereof.

In a further embodiment, L^9 is N-R^9, wherein R^9 is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, acyl, or aryl. In another, L^{10} is oxygen. In an embodiment, R^9 is hydrogen. In another, G is substituted or unsubstituted phenyl or heteroaryl. In a further embodiment, G is cycloalkenyl, e.g., cyclohexenyl. In one embodiment, L^1, L^2, L^3, and L^4 are each substituted or unsubstituted carbon and L^5, L^6, and L^8 are each nitrogen. L^7 may be substituted carbon, e.g., substituted with a thioether moiety. In another embodiment, L^9 and L^{10} are each nitrogen. In another embodiment, the invention pertains to compounds of formula (III), wherein L^9 is nitrogen, L^{10} is oxygen, L^1-L^8 are each C-H, the dotted line represents a double bond and and where G is not hydrogen or methyl.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (X):



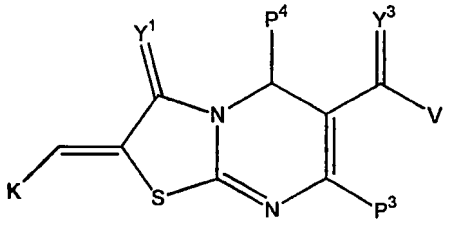
(X)

wherein

Y^1 and Y^2 are each oxygen, sulfur, or substituted or unsubstituted carbon;

J^1, J^2, J^3, and J^4 are each oxygen, nitrogen, or optionally substituted carbon, and pharmaceutically acceptable salts thereof.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (IV):



(IV)

wherein

Y^1 and Y^2 are each oxygen or sulfur;

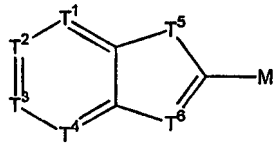
J is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, cyano, nitro, amino, or halogen;

V is substituted or unsubstituted alkyl, alkenyl, alkynyl, alkoxy, alkylamino, or alkylthio;

P and K are each independently substituted or unsubstituted aryl, and pharmaceutically acceptable salts thereof.

In a further embodiment, Y^1 and Y^3 are each oxygen, V is alkoxy and J is lower alkyl. In another embodiment, P is substituted or unsubstituted phenyl. K may be substituted or unsubstituted heteroaryl.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (V):



(V)

wherein

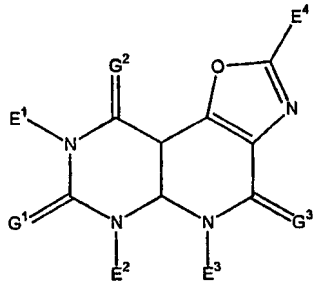
T^1, T^2, T^3, T^4, T^5, and T^6 are each independently substituted or unsubstituted carbon, oxygen, substituted or unsubstituted nitrogen, or sulfur;

M is hydrogen, alkyl, alkenyl, alkynyl, heterocyclic or aryl, or pharmaceutically acceptable salts thereof.

In a further embodiment, T^5 is N-W or C-W, wherein W is alkyl, alkenyl, alkynyl, aryl, heterocyclic, acyl, hydroxy, alkoxy, alkthio, amino, nitro, halogen, or hydrogen. In another further embodiment, T^6 is N.

In a further embodiment, M is substituted or unsubstituted aryl. W may be substituted or unsubstituted alkyl. In another embodiment, T^1, T^2, T^3 and T^4 are each substituted or unsubstituted carbon. In a further embodiment, at least one of T^1, T^2, T^3, and T^4 is nitrogen, and the remaining T moieties are substituted or unsubstituted carbon.

In another embodiment, the transcription factor modulating compound (e.g., an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, etc.) is of the formula (VI):



(VI)

wherein

G^1, G^2, and G^3 are each independently O, S, substituted or unsubstituted nitrogen, or substituted or unsubstituted carbon;

E^1, E^2, and E^3 are each independently hydrogen, alkyl, alkenyl, alkynyl, aryl, aralkyl, or acyl; and

E^4 is alkyl, alkenyl, alkynyl, aryl, halogen, cyano, amino, nitro, or acyl, and pharmaceutically acceptable salts thereof.

In a further embodiment, G^1, G^2 and G^3 are each oxygen.

Other transcription factor modulating compounds are shown in Table 3. The invention pertains to each of these compounds, methods (both therapeutic and otherwise) using each of the compounds, and compositions comprising at least one of the compounds of Table 4, Table 5, or of formulae (I), (II), (III), (IV), (V), (VI), (VII), (VIII), (IX) or (X).

The invention also pertains to each of the following compounds:

2-(4-isopropylphenyl)-4H-chromen-4-one;

2-(3,4-Dihydroxy-phenyl)-3,5,7-trihydroxy-chromen-4-one

N-isopropyl-2-[(4-methyl-5-quinolin-6-yl-4H-1,2,4-triazol-3-yl)thio]acetamide;

4-hydroxy-6-methyl-5,6-dihydro-2H-pyrano[3,2-c]quinoline-2,5-dione;

5,7-Dihydroxy-2-(4-hydroxy-phenyl)-chromen-4-one;

2-[4-(dimethylamino)phenyl]-4H-chromen-4-one;

1-(benzyloxy)-2-phenyl-1H-imidazo[4,5-b]pyridine;

2-(benzylthio)-4-phenyl-5-(1-phenyl-1H-1,2,3,4-tetraazol-5-yl)pyrimidine;

6-fluoro-2-phenyl-4H-chromen-4-one;

7-methoxy-2-phenyl-4H-chromen-4-one;

4-(1,3-dioxo-1,3-dihydro-2H-inden-2-yliden)-2-phenyl-6-(2-pyridinyl)tetrahydropyrrolo[3,4-c]pyrrole-1,3(2H,3aH)-dione;

2-(2-Hydroxy-3-oxo-5-p-tolyl-2,3-dihydro-furan-2-yl)-malonamic acid ethyl ester;

2-[(6-nitro-2-phenyl-1H-1,3-benzimidazol-1-yl)oxy]acetic acid;

2-(4-fluorophenyl)-4H-chromen-4-one;

1-methoxy-2-(4-methylphenyl)-1H-imidazo[4,5-b]pyridine;

6-(5-Iodo-furan-2-yl)-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

6-(4-Ethoxy-phenyl)-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

3-Methylsulfanyl-6-(5-nitro-furan-2-yl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

3-Methylsulfanyl-6-[5-(4-nitro-phenyl)-furan-2-yl]-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo [a,c] cycloheptene;

4-(3-Ethylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo [a,c] cyclohepten-6-yl)-benzene-1,2-diol;

6-(4-Benzyloxy-phenyl)-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

6-Benzo[1,3]dioxol-5-yl-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

3-Butylsulfanyl-6-(2,4-dimethoxy-phenyl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo[a,c] cycloheptene;

6-(4-Allyloxy-phenyl)-3-butylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo[a,c]cycloheptene;

3-Butylsulfanyl-6-(4-ethoxy-phenyl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c] cycloheptene;

6-(4-Methoxy-phenyl)-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

6-[5-(3-Nitro-phenyl)-furan-2-yl]-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo[a,c] cycloheptene;

2-(3-Phenyl-1H-pyrazol-4-ylmethylene)-benzo[4,5] imidazo[2,1-b]thiazol-3-one;

2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-5-(2-methoxy-naphthalen-1-yl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

5-(4-Dimethylamino-phenyl)-7-methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro -5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester;

5-Benzo[1,3]dioxol-5-yl-7-methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro -5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester;

5-(3,4-Dimethoxy-phenyl)-7-methyl-2-[5-(2-methyl-4-nitro -phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

7-Methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-ylmethylene]-5-(4-methyl sulfanyl-phenyl)-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester;

2-[5-(4-Carboxy-phenyl)-furan-2-ylmethylene]-5-(2-methoxy-naphthalen-1-yl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

5-Benzo[1,3]dioxol-5-yl-2-[5-(4-ethoxycarbonyl-phenyl)-furan-2-ylmethylene]-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

7-Methyl-3-oxo-5-phenyl-2-[5-(3-trifluoromethyl-phenyl)-furan-2-ylmethylene]-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

7-Methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-5-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-5-(4-dimethylamino-phenyl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

5-(4-Dimethylamino-phenyl)-7-methyl-2-[5-(4-methyl-3-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-7-methyl-5-(4-methylsulfanyl-phenyl)-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

[1,2]Naphthoquinone 1-[O-(6-oxo-6H-anthra[1,9-cd] isoxazol-5-yl)-oxime];

3-Acetyl-2,5,7-triphenyl-1H-1,3a,4,8-tetraaza-7a-azonia-cyclopenta[a]indene;

1-Amino-3-benzo[1,3]dioxol-5-yl-benzo[4,5]imidazo[1,2-a] pyridine-2,4-dicarbonitrile;

2-[2-(5-Furan-2-yl-4-phenyl-4H-[1,2,4]triazol-3-yl sulfanyl)-acetylamino]-benzoic acid methyl ester;

6,7-Dimethyl-2-(3-phenyl-1H-pyrazol-4-ylmethylene)-benzo[4,5]imidazo[2,1-b] thiazol-3-one;

2-(5-Benzo[1,2,5]oxadiazol-5-yl-4-methyl-4H-[1,2,4] triazol-3-ylsulfanyl)-N-(3-methylsulfanyl-phenyl)-acetamide;

4-(1,3-Dioxo-indan-2-ylidene)-2-phenyl-6-pyridin-2-yl-tetrahydro-pyrrolo[3,4-c] pyrrole-1,3-dione;

6-Nitro-2-phenyl-1-(3-trifluoromethyl-benzyloxy)-1H-benzoimidazole;

(6-Nitro-2-phenyl-benzoimidazol-1-yloxy)-acetic acid;

1-Benzyloxy-6-nitro-2-phenyl-1H-benzoimidazole;

1-(4-Methyl-benzyloxy)-6-nitro-2-phenyl-1H-benzoimidazole;

6,8-Dimethyl-2-(4-nitro-phenyl)-5-phenyl-5H,6H-1-oxa-3,5,6,8-tetraaza-cyclopenta[a]naphthalene-4,7,9-trione;

6,8-Dimethyl-5-phenyl-2-p-tolyl-5H,6H-1-oxa-3,5,6,8-tetraaza-cyclopenta [a]naphthalene-4,7,9-trione;

2-[3-(4-Fluoro-phenyl)-1-phenyl-1H-pyrazol-4-yl methylene]-benzo [4,5] imidazo[2,1-b]thiazol-3-one;

Cobalt 5,10,15,20-Tetra-pyridin-4-yl-porphyrine;

2-[3-(4-Fluoro-phenyl)-1-phenyl-1H-pyrazol-4-ylmethylene]-5-methyl-6-vinyl-imidazo[2,1-b]thiazol-3-one;

Cobalt 5,10,15,20-Tetra-pyridin-3-yl-porphyrine;

Zinc 5,10,15,20-Tetra-pyridin-4-yl-porphyrine;

2-(4-hydroxyphenyl)-4H-chromen-4-one, and pharmaceutically acceptable salts thereof.

In a further embodiment, the transcription factor modulating compound is not apigenin. In another

The term "alkyl" includes saturated aliphatic groups, including straight-chain alkyl groups (*e.g.*, methyl, ethyl, propyl, butyl, pentyl, hexyl, heptyl, octyl, nonyl, decyl, etc.), branched-chain alkyl groups (isopropyl, tert-butyl, isobutyl, etc.), cycloalkyl (alicyclic) groups (cyclopropyl, cyclopentyl, cyclohexyl, cycloheptyl, cyclooctyl), alkyl substituted cycloalkyl groups, and cycloalkyl substituted alkyl groups. The term alkyl further includes alkyl groups, which can further include oxygen, nitrogen, sulfur or

phosphorous atoms replacing one or more carbons of the hydrocarbon backbone. In certain embodiments, a straight chain or branched chain alkyl has 6 or fewer carbon atoms in its backbone (*e.g.*, C_1-C_6 for straight chain, C_3-C_6 for branched chain), and more preferably 4 or fewer. Likewise, preferred cycloalkyls have from 3-8 carbon atoms in their ring structure, and more preferably have 5 or 6 carbons in the ring structure. The term C_1-C_6 includes alkyl groups containing 1 to 6 carbon atoms.

Moreover, the term alkyl includes both "unsubstituted alkyls" and "substituted alkyls", the latter of which refers to alkyl moieties having substituents replacing a hydrogen on one or more carbons of the hydrocarbon backbone. Such substituents can include, for example, alkenyl, alkynyl, halogen, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, arylcarbonyl, alkoxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moiety. Cycloalkyls can be further substituted, *e.g.*, with the substituents described above. An "alkylaryl" or an "arylalkyl" moiety is an alkyl substituted with an aryl (*e.g.*, phenylmethyl (benzyl)). The term "alkyl" also includes the side chains of natural and unnatural amino acids.

The term "aryl" includes groups, including 5- and 6-membered single-ring aromatic groups that may include from zero to four heteroatoms, for example, benzene, phenyl, pyrrole, furan, thiophene, thiazole, isothiaozole, imidazole, triazole, tetrazole, pyrazole, oxazole, isooxazole, pyridine, pyrazine, pyridazine, and pyrimidine, and the like. Furthermore, the term "aryl" includes multicyclic aryl groups, *e.g.*, tricyclic, bicyclic, *e.g.*, naphthalene, benzoxazole, benzodioxazole, benzothiazole, benzoimidazole, benzothiophene, methylenedioxyphenyl, quinoline, isoquinoline, napthridine, indole, benzofuran, purine, benzofuran, deazapurine, or indolizine. Those aryl groups having heteroatoms in the ring structure may also be referred to as "aryl heterocycles", "heterocycles," "heteroaryls" or "heteroaromatics". The aromatic ring can be substituted at one or more ring positions with such substituents as described above, as for example, halogen, hydroxyl, alkoxy, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, alkylaminoacarbonyl, arylalkyl aminocarbonyl, alkenylaminocarbonyl, alkylcarbonyl, arylcarbonyl, arylalkylcarbonyl, alkenylcarbonyl, alkoxycarbonyl, aminocarbonyl,

alkylthiocarbonyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moiety. Aryl groups can also be fused or bridged with alicyclic or heterocyclic rings which are not aromatic so as to form a polycycle (*e.g.*, tetralin). The term "aryl" also includes multicyclic aryl groups such as porphrins, phthalocyanines, etc.

The term "alkenyl" includes unsaturated aliphatic groups analogous in length and possible substitution to the alkyls described above, but that contain at least one double bond.

For example, the term "alkenyl" includes straight-chain alkenyl groups (*e.g.*, ethylenyl, propenyl, butenyl, pentenyl, hexenyl, heptenyl, octenyl, nonenyl, decenyl, etc.), branched-chain alkenyl groups, cycloalkenyl (alicyclic) groups (cyclopropenyl, cyclopentenyl, cyclohexenyl, cycloheptenyl, cyclooctenyl), alkyl or alkenyl substituted cycloalkenyl groups, and cycloalkyl or cycloalkenyl substituted alkenyl groups. The term alkenyl further includes alkenyl groups which include oxygen, nitrogen, sulfur or phosphorous atoms replacing one or more carbons of the hydrocarbon backbone. In certain embodiments, a straight chain or branched chain alkenyl group has 6 or fewer carbon atoms in its backbone (*e.g.*, C_2-C_6 for straight chain, C_3-C_6 for branched chain). Likewise, cycloalkenyl groups may have from 3-8 carbon atoms in their ring structure, and more preferably have 5 or 6 carbons in the ring structure. The term C_2-C_6 includes alkenyl groups containing 2 to 6 carbon atoms.

Moreover, the term alkenyl includes both "unsubstituted alkenyls" and "substituted alkenyls", the latter of which refers to alkenyl moieties having substituents replacing a hydrogen on one or more carbons of the hydrocarbon backbone. Such substituents can include, for example, alkyl groups, alkynyl groups, halogens, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, arylcarbonyl, alkoxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moiety.

The term "alkynyl" includes unsaturated aliphatic groups analogous in length and possible substitution to the alkyls described above, but which contain at least one triple bond.

For example, the term "alkynyl" includes straight-chain alkynyl groups (*e.g.*, ethynyl, propynyl, butynyl, pentynyl, hexynyl, heptynyl, octynyl, nonynyl, decynyl, etc.), branched-chain alkynyl groups, and cycloalkyl or cycloalkenyl substituted alkynyl groups. The term alkynyl further includes alkynyl groups which include oxygen, nitrogen, sulfur or phosphorous atoms replacing one or more carbons of the hydrocarbon backbone. In certain embodiments, a straight chain or branched chain alkynyl group has 6 or fewer carbon atoms in its backbone (*e.g.*, C_2-C_6 for straight chain, C_3-C_6 for branched chain). The term C_2-C_6 includes alkynyl groups containing 2 to 6 carbon atoms.

Moreover, the term alkynyl includes both "unsubstituted alkynyls" and "substituted alkynyls", the latter of which refers to alkynyl moieties having substituents replacing a hydrogen on one or more carbons of the hydrocarbon backbone. Such substituents can include, for example, alkyl groups, alkynyl groups, halogens, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, arylcarbonyl, alkoxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moiety.

Unless the number of carbons is otherwise specified, "lower alkyl" as used herein means an alkyl group, as defined above, but having from one to five carbon atoms in its backbone structure. "Lower alkenyl" and "lower alkynyl" have chain lengths of, for example, 2-5 carbon atoms.

The term "acyl" includes compounds and moieties which contain the acyl radical (CH_3CO-) or a carbonyl group. The term "substituted acyl" includes acyl groups where one or more of the hydrogen atoms are replaced by for example, alkyl groups, alkynyl groups, halogens, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, arylcarbonyl, alkoxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including

alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moiety.

The term "acylamino" includes moieties wherein an acyl moiety is bonded to an amino group. For example, the term includes alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido groups.

The term "aroyl" includes compounds and moieties with an aryl or heteroaromatic moiety bound to a carbonyl group. Examples of aroyl groups include phenylcarboxy, naphthyl carboxy, etc.

The terms "alkoxyalkyl", "alkylaminoalkyl" and "thioalkoxyalkyl" include alkyl groups, as described above, which further include oxygen, nitrogen or sulfur atoms replacing one or more carbons of the hydrocarbon backbone, *e.g.*, oxygen, nitrogen or sulfur atoms.

The term "alkoxy" includes substituted and unsubstituted alkyl, alkenyl, and alkynyl groups covalently linked to an oxygen atom. Examples of alkoxy groups include methoxy, ethoxy, isopropyloxy, propoxy, butoxy, and pentoxy groups. Examples of substituted alkoxy groups include halogenated alkoxy groups. The alkoxy groups can be substituted with groups such as alkenyl, alkynyl, halogen, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, arylcarbonyl, alkoxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkylaryl, or an aromatic or heteroaromatic moieties. Examples of halogen substituted alkoxy groups include, but are not limited to, fluoromethoxy, difluoromethoxy, trifluoromethoxy, chloromethoxy, dichloromethoxy, trichloromethoxy, etc.

The term "amine" or "amino" includes compounds where a nitrogen atom is covalently bonded to at least one carbon or heteroatom. The term "alkyl amino" includes groups and compounds wherein the nitrogen is bound to at least one additional alkyl group. The term "dialkyl amino" includes groups wherein the nitrogen atom is bound to at least two additional alkyl groups. The term "arylamino" and "diarylamino" include groups wherein the nitrogen is bound to at least one or two aryl groups, respectively. The term "alkylarylamino," "alkylaminoaryl" or "arylaminoalkyl" refers

to an amino group which is bound to at least one alkyl group and at least one aryl group. The term "alkaminoalkyl" refers to an alkyl, alkenyl, or alkynyl group bound to a nitrogen atom which is also bound to an alkyl group.

The term "amide" or "aminocarboxy" includes compounds or moieties which contain a nitrogen atom which is bound to the carbon of a carbonyl or a thiocarbonyl group. The term includes "alkaminocarboxy" groups which include alkyl, alkenyl, or alkynyl groups bound to an amino group bound to a carboxy group. It includes arylaminocarboxy groups which include aryl or heteroaryl moieties bound to an amino group which is bound to the carbon of a carbonyl or thiocarbonyl group. The terms "alkylaminocarboxy," "alkenylaminocarboxy," "alkynylaminocarboxy," and "arylaminocarboxy" include moieties wherein alkyl, alkenyl, alkynyl and aryl moieties, respectively, are bound to a nitrogen atom which is in turn bound to the carbon of a carbonyl group.

The term "carbonyl" or "carboxy" includes compounds and moieties which contain a carbon connected with a double bond to an oxygen atom. Examples of moieties which contain a carbonyl include aldehydes, ketones, carboxylic acids, amides, esters, anhydrides, etc.

The term "thiocarbonyl" or "thiocarboxy" includes compounds and moieties which contain a carbon connected with a double bond to a sulfur atom.

The term "ether" includes compounds or moieties which contain an oxygen bonded to two different carbon atoms or heteroatoms. For example, the term includes "alkoxyalkyl" which refers to an alkyl, alkenyl, or alkynyl group covalently bonded to an oxygen atom which is covalently bonded to another alkyl group.

The term "ester" includes compounds and moieties which contain a carbon or a heteroatom bound to an oxygen atom which is bonded to the carbon of a carbonyl group. The term "ester" includes alkoxycarboxy groups such as methoxycarbonyl, ethoxycarbonyl, propoxycarbonyl, butoxycarbonyl, pentoxycarbonyl, etc. The alkyl, alkenyl, or alkynyl groups are as defined above.

The term "thioether" includes compounds and moieties which contain a sulfur atom bonded to two different carbon or hetero atoms. Examples of thioethers include, but are not limited to alkthioalkyls, alkthioalkenyls, and alkthioalkynyls. The term "alkthioalkyls" include compounds with an alkyl, alkenyl, or alkynyl group bonded to a sulfur atom which is bonded to an alkyl group. Similarly, the term "alkthioalkenyls" and alkthioalkynyls" refer to compounds or moieties wherein an alkyl, alkenyl, or alkynyl group is bonded to a sulfur atom which is covalently bonded to an alkynyl group.

The term "hydroxy" or "hydroxyl" includes groups with an –OH or $-O^-$.

The term "halogen" includes fluorine, bromine, chlorine, iodine, etc. The term "perhalogenated" generally refers to a moiety wherein all hydrogens are replaced by halogen atoms.

The terms "polycyclyl" or "polycyclic radical" refer to two or more cyclic rings (*e.g.*, cycloalkyls, cycloalkenyls, cycloalkynyls, aryls and/or heterocyclyls) in which two or more carbons are common to two adjoining rings, *e.g.*, the rings are "fused rings". Rings that are joined through non-adjacent atoms are termed "bridged" rings. Each of the rings of the polycycle can be substituted with such substituents as described above, as for example, halogen, hydroxyl, alkylcarbonyloxy, arylcarbonyloxy, alkoxycarbonyloxy, aryloxycarbonyloxy, carboxylate, alkylcarbonyl, alkoxycarbonyl, alkylaminoacarbonyl, arylalkylaminocarbonyl, alkenylaminocarbonyl, alkylcarbonyl, arylcarbonyl, arylalkyl carbonyl, alkenylcarbonyl, aminocarbonyl, alkylthiocarbonyl, alkoxyl, phosphate, phosphonato, phosphinato, cyano, amino (including alkyl amino, dialkylamino, arylamino, diarylamino, and alkylarylamino), acylamino (including alkylcarbonylamino, arylcarbonylamino, carbamoyl and ureido), amidino, imino, sulfhydryl, alkylthio, arylthio, thiocarboxylate, sulfates, alkylsulfinyl, sulfonato, sulfamoyl, sulfonamido, nitro, trifluoromethyl, cyano, azido, heterocyclyl, alkyl, alkylaryl, or an aromatic or heteroaromatic moiety.

The term "heteroatom" includes atoms of any element other than carbon or hydrogen. Preferred heteroatoms are nitrogen, oxygen, sulfur and phosphorus.

It will be noted that the structure of some of the compounds of this invention includes asymmetric carbon atoms. It is to be understood accordingly that the isomers arising from such asymmetry (*e.g.*, all enantiomers and diastereomers) are included within the scope of this invention, unless indicated otherwise. Such isomers can be obtained in substantially pure form by classical separation techniques and by stereochemically controlled synthesis. Furthermore, the structures and other compounds and moieties discussed in this application also include all tautomers thereof.

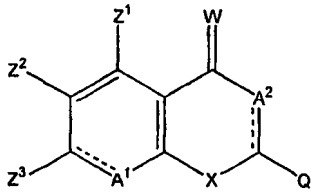
Bonds represented by " ------ " in a structural formula mean that the bond may be either a single or a double bond.

IX. Formulations Comprising Transcription factor Modulating Compounds

The invention provides compositions which include a therapeutically-effective amount or dose of a transcription factor modulating compound and/or a compound identified in any of the instant assays and one or more carriers (*e.g.*, pharmaceutically acceptable additives and/or diluents). The pharmaceutical compositions of the invention may comprise any compound described in this application as a transcription factor modulating compound, an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, a MarA family inhibiting

compound, a MarA inhibiting compound, compounds of formulae (I), (II), (III), (IV), (V), (VI), (VII), (VIII), (IX), (X), Table 4, Table 5, scaffold, etc. Each of these compounds may be used alone of in combination as a part of a pharmaceutical composition of the invention. Furthermore, a composition can also include a second antimicrobial agent, *e.g.*, an antibiotic.

The invention pertains to pharmaceutical compositions comprising an effective amount of a transcription factor modulating compound (e.g., a MarA family polypeptide modulating compound or an AraC family polypeptide modulating compound), and a pharmaceutically acceptable carrier. In one embodiment, the transcription factor modulating compound is of the formula (II):



(II)

wherein

W is O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^4, O, or S;

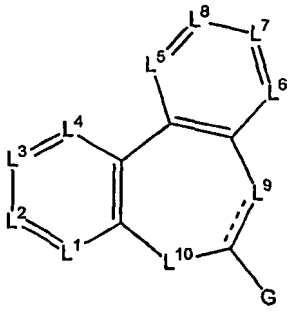
A^2 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, Z^4 and Z^5 are each independently hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, or cyano;

Z^3 is hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, nitro, cyano, carbonyl, or thiocarbonyl;

Q is an aromatic or heterocyclic moiety, and pharmaceutically acceptable salts thereof.

In another embodiment, the pharmaceutical compositions of the invention include an effective amount of a transcription factor modulating compound of the formula (III):



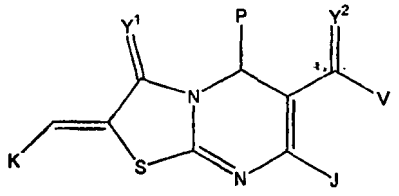
(III)

wherein

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen; and

L^1, L^2, L^3, L^4, L^5, L^6, L^7, L^8, L^9, and L^{10} are each independently oxygen, substituted or unsubstituted nitrogen, sulfur and or substituted or unsubstituted carbon, and pharmaceutically acceptable salts thereof.

In yet another embodiment, the pharmaceutical compositions of the invention include a pharmaceutically acceptable carrier (optional) and an effective amount of a transcription factor modulating compound of the formula (IV):



(IV)

wherein

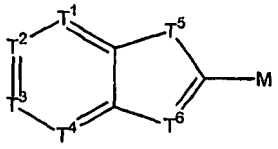
Y^1 and Y^2 are each oxygen or sulfur;

J is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, cyano, nitro, amino, or halogen;

V is substituted or unsubstituted alkyl, alkenyl, alkynyl, alkoxy, alkylamino, or alkylthio;

P and K are each independently substituted or unsubstituted aryl, and pharmaceutically acceptable salts thereof.

In yet another embodiment, the pharmaceutical compositions of the invention include a pharmaceutically acceptable carrier (optional) and an effective amount of a transcription factor modulating compound of the formula (V):



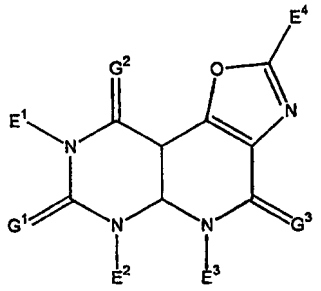
(V)

wherein

T^1, T^2, T^3, T^4, T^5, and T^6 are each independently substituted or unsubstituted carbon, oxygen, substituted or unsubstituted nitrogen, or sulfur;

M is hydrogen, alkyl, alkenyl, alkynyl, or aryl, or pharmaceutically acceptable salts thereof.

In yet another embodiment, the pharmaceutical compositions of the invention include a pharmaceutically acceptable carrier (optional) and an effective amount of a transcription factor modulating compound of the formula (VI):



(VI)

wherein

G^1, G^2, and G^3 are each independently O, S, substituted or unsubstituted nitrogen, or substituted or unsubstituted carbon;

E^1, E^2, and E^3 are each independently hydrogen, alkyl, alkenyl, alkynyl, aryl, aralkyl, or acyl; and

E^4 is alkyl, alkenyl, alkynyl, aryl, halogen, cyano, amino, nitro, or acyl, and pharmaceutically acceptable salts thereof.

In yet another further embodiment, the pharmaceutical compositions of the invention comprise an effective amount of a transcription factor modulating compound listed below or found in Table 4 or Table 5:

2-(4-isopropylphenyl)-4H-chromen-4-one;

2-(3,4-Dihydroxy-phenyl)-3,5,7-trihydroxy-chromen-4-one

N-isopropyl-2-[(4-methyl-5-quinolin-6-yl-4H-1,2,4-triazol-3-yl)thio]acetamide;

4-hydroxy-6-methyl-5,6-dihydro-2H-pyrano[3,2-c]quinoline-2,5-dione;

5,7-Dihydroxy-2-(4-hydroxy-phenyl)-chromen-4-one;

2-[4-(dimethylamino)phenyl]-4H-chromen-4-one;

1-(benzyloxy)-2-phenyl-1H-imidazo[4,5-b]pyridine;

2-(benzylthio)-4-phenyl-5-(1-phenyl-1H-1,2,3,4-tetraazol-5-yl)pyrimidine;

6-fluoro-2-phenyl-4H-chromen-4-one;

7-methoxy-2-phenyl-4H-chromen-4-one;

4-(1,3-dioxo-1,3-dihydro-2H-inden-2-yliden)-2-phenyl-6-(2-pyridinyl)tetrahydropyrrolo[3,4-c]pyrrole-1,3(2H,3aH)-dione;

2-(2-Hydroxy-3-oxo-5-p-tolyl-2,3-dihydro-furan-2-yl)-malonamic acid ethyl ester;

2-[(6-nitro-2-phenyl-1H-1,3-benzimidazol-1-yl)oxy]acetic acid;

2-(4-fluorophenyl)-4H-chromen-4-one;

1-methoxy-2-(4-methylphenyl)-1H-imidazo[4,5-b]pyridine;

6-(5-Iodo-furan-2-yl)-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

6-(4-Ethoxy-phenyl)-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

3-Methylsulfanyl-6-(5-nitro-furan-2-yl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

3-Methylsulfanyl-6-[5-(4-nitro-phenyl)-furan-2-yl]-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo [a,c] cycloheptene;

4-(3-Ethylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo [a,c] cyclohepten-6-yl)-benzene-1,2-diol;

6-(4-Benzyloxy-phenyl)-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

6-Benzo[1,3]dioxol-5-yl-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

3-Butylsulfanyl-6-(2,4-dimethoxy-phenyl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo[a,c] cycloheptene;

6-(4-Allyloxy-phenyl)-3-butylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo[a,c]cycloheptene;

3-Butylsulfanyl-6-(4-ethoxy-phenyl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c] cycloheptene;

6-(4-Methoxy-phenyl)-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene;

6-[5-(3-Nitro-phenyl)-furan-2-yl]-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo[a,c] cycloheptene;

2-(3-Phenyl-1H-pyrazol-4-ylmethylene)-benzo[4,5] imidazo[2,1-b]thiazol-3-one;

2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-5-(2-methoxy-naphthalen-1-yl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

5-(4-Dimethylamino-phenyl)-7-methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro -5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester;

5-Benzo[1,3]dioxol-5-yl-7-methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro -5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester;

5-(3,4-Dimethoxy-phenyl)-7-methyl-2-[5-(2-methyl-4-nitro -phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

7-Methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-ylmethylene]-5-(4-methyl sulfanyl-phenyl)-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester;

2-[5-(4-Carboxy-phenyl)-furan-2-ylmethylene]-5-(2-methoxy-naphthalen-1-yl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

5-Benzo[1,3]dioxol-5-yl-2-[5-(4-ethoxycarbonyl-phenyl)-furan-2-ylmethylene]-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

7-Methyl-3-oxo-5-phenyl-2-[5-(3-trifluoromethyl-phenyl)-furan-2-ylmethylene]-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

7-Methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-5-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-5-(4-dimethylamino-phenyl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

5-(4-Dimethylamino-phenyl)-7-methyl-2-[5-(4-methyl-3-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-7-methyl-5-(4-methylsulfanyl-phenyl)-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester;

[1,2]Naphthoquinone 1-[O-(6-oxo-6H-anthra[1,9-cd] isoxazol-5-yl)-oxime];

3-Acetyl-2,5,7-triphenyl-1H-1,3a,4,8-tetraaza-7a-azonia-cyclopenta[a]indene;

1-Amino-3-benzo[1,3]dioxol-5-yl-benzo[4,5]imidazo[1,2-a] pyridine-2,4-dicarbonitrile;

2-[2-(5-Furan-2-yl-4-phenyl-4H-[1,2,4]triazol-3-yl sulfanyl)-acetylamino]-benzoic acid methyl ester;

6,7-Dimethyl-2-(3-phenyl-1H-pyrazol-4-ylmethylene)-benzo[4,5]imidazo[2,1-b] thiazol-3-one;

2-(5-Benzo[1,2,5]oxadiazol-5-yl-4-methyl-4H-[1,2,4] triazol-3-ylsulfanyl)-N-(3-methylsulfanyl-phenyl)-acetamide;

4-(1,3-Dioxo-indan-2-ylidene)-2-phenyl-6-pyridin-2-yl-tetrahydro-pyrrolo[3,4-c] pyrrole-1,3-dione;

6-Nitro-2-phenyl-1-(3-trifluoromethyl-benzyloxy)-1H-benzoimidazole;

(6-Nitro-2-phenyl-benzoimidazol-1-yloxy)-acetic acid;

1-Benzyloxy-6-nitro-2-phenyl-1H-benzoimidazole;

1-(4-Methyl-benzyloxy)-6-nitro-2-phenyl-1H-benzoimidazole;

6,8-Dimethyl-2-(4-nitro-phenyl)-5-phenyl-5H,6H-1-oxa-3,5,6,8-tetraaza-cyclopenta[a]naphthalene-4,7,9-trione;

6,8-Dimethyl-5-phenyl-2-p-tolyl-5H,6H-1-oxa-3,5,6,8-tetraaza-cyclopenta [a]naphthalene-4,7,9-trione;

2-[3-(4-Fluoro-phenyl)-1-phenyl-1H-pyrazol-4-yl methylene]-benzo [4,5] imidazo[2,1-b]thiazol-3-one;

Cobalt 5,10,15,20-Tetra-pyridin-4-yl-porphyrine;

2-[3-(4-Fluoro-phenyl)-1-phenyl-1H-pyrazol-4-ylmethylene]-5-methyl-6-vinyl-imidazo[2,1-b]thiazol-3-one;

Cobalt 5,10,15,20-Tetra-pyridin-3-yl-porphyrine;

Zinc 5,10,15,20-Tetra-pyridin-4-yl-porphyrine;

2-(4-hydroxyphenyl)-4H-chromen-4-one, and pharmaceutically acceptable salts thereof.

In another embodiment, the method for preventing a bacterial associated state in a subject, comprising administering to the subject an effective amount of a transcription factor modulating compound, such that the bacterial associated state is prevented.

The term "subject" includes plants and animals (*e.g.*, vertebrates, amphibians, fish, mammals, *e.g.*, cats, dogs, horses, pigs, cows, sheep, rodents, rabbits, squirrels, bears, primates (*e.g.*, chimpanzees, gorillas, and humans) which are capable of suffering from a bacterial associated disorder. The term "subject" also comprises immunocompromised subjects, who may be at a higher risk for infection.

The term "preventing" the administration of an effective amount of the transcription factor modulating compound to prevent a bacterial associated state from occurring.

The term "bacterial associated state" includes states characterized by the presence of bacteria which can be prevented by administering the transcription factor modulating compounds of the invention. The term includes biofilm associated states and other infections or the undesirable presence of a bacteria on or in a subject.

As described in detail below, the pharmaceutical compositions can be formulated for administration in solid or liquid form, including those adapted for the following: (1) oral administration, for example, aqueous or non-aqueous solutions or suspensions, tablets, boluses, powders, granules, pastes; (2) parental administration, for example, by subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution or suspension; (3) topical application, for example, as a cream, ointment or spray applied to the skin; (4) intravaginally or intrarectally, for example, as a pessary, cream, foam, or suppository; or (5) aerosol, for example, as an aqueous aerosol, liposomal preparation or solid particles containing the compound.

The phrase "pharmaceutically-acceptable carrier" as used herein means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting the antiinfective agents or compounds of the invention from one organ, or portion of the body, to another organ, or portion of the body without affecting its biological effect. Each carrier should be "acceptable" in the sense of being compatible with the other ingredients of the composition and not injurious to the subject. Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) phosphate buffer solutions; and (21) other non-toxic compatible substances employed in pharmaceutical compositions. Proper fluidity can be maintained, for example, by the use of coating materials, such as lecithin, by the maintenance of the required particle size in the case of dispersions, and by the use of surfactants.

These compositions may also contain adjuvants such as preservatives, wetting agents, emulsifying agents and dispersing agents. Prevention of the action of microbes may be ensured by the inclusion of various antibacterial and antifungal agents, for example, paraben, chlorobutanol, phenol sorbic acid, and the like. It may also be desirable to include isotonic agents, such as sugars, sodium chloride, and the like into the compositions. In addition, prolonged absorption of the injectable pharmaceutical

form may be brought about by the inclusion of agents which delay absorption such as aluminum monostearate and gelatin.

In some cases, in order to prolong the effect of a drug, it is desirable to slow the absorption of the drug from subcutaneous or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material having poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution which, in turn, may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally-administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle.

Pharmaceutical compositions of the present invention may be administered to epithelial surfaces of the body orally, parenterally, topically, rectally, nasally, intravaginally, intracisternally. They are of course given by forms suitable for each administration route. For example, they are administered in tablets or capsule form, by injection, inhalation, eye lotion, ointment, etc., administration by injection, infusion or inhalation; topical by lotion or ointment; and rectal or vaginal suppositories.

The phrases "parenteral administration" and "administered parenterally" as used herein mean modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, subcapsular, subarachnoid, intraspinal and intrasternal injection and infusion.

The phrases "systemic administration," "administered systemically," "peripheral administration" and "administered peripherally" as used herein mean the administration of a sucrose octasulfate and/or an antibacterial, drug or other material other than directly into the central nervous system, such that it enters the subject's system and, thus, is subject to metabolism and other like processes, for example, subcutaneous administration.

In some methods, the compositions of the invention can be topically administered to any epithelial surface. An "epithelial surface" according to this invention is defined as an area of tissue that covers external surfaces of a body, or which lines hollow structures including, but not limited to, cutaneous and mucosal surfaces. Such epithelial surfaces include oral, pharyngeal, esophageal, pulmonary, ocular, aural, nasal, buccal, lingual, vaginal, cervical, genitourinary, alimentary, and anorectal surfaces.

Compositions can be formulated in a variety of conventional forms employed for topical administration. These include, for example, semi-solid and liquid dosage forms, such as liquid solutions or suspensions, suppositories, douches, enemas,

gels, creams, emulsions, lotions, slurries, powders, sprays, lipsticks, foams, pastes, toothpastes, ointments, salves, balms, douches, drops, troches, chewing gums, lozenges, mouthwashes, rinses.

Conventionally used carriers for topical applications include pectin, gelatin and derivatives thereof, polylactic acid or polyglycolic acid polymers or copolymers thereof, cellulose derivatives such as methyl cellulose, carboxymethyl cellulose, or oxidized cellulose, guar gum, acacia gum, karaya gum, tragacanth gum, bentonite, agar, carbomer, bladderwrack, ceratonia, dextran and derivatives thereof, ghatti gum, hectorite, ispaghula husk, polyvinypyrrolidone, silica and derivatives thereof, xanthan gum, kaolin, talc, starch and derivatives thereof, paraffin, water, vegetable and animal oils, polyethylene, polyethylene oxide, polyethylene glycol, polypropylene glycol, glycerol, ethanol, propanol, propylene glycol (glycols, alcohols), fixed oils, sodium, potassium, aluminum, magnesium or calcium salts (such as chloride, carbonate, bicarbonate, citrate, gluconate, lactate, acetate, gluceptate or tartrate).

Such compositions can be particularly useful, for example, for treatment or prevention of an unwanted cell, *e.g.*, vaginal *Neisseria gonorrhoeae*, or infections of the oral cavity, including cold sores, infections of eye, the skin, or the lower intestinal tract. Standard composition strategies for topical agents can be applied to the antiinfective compounds or a pharmaceutically acceptable salt thereof in order to enhance the persistence and residence time of the drug, and to improve the prophylactic efficacy achieved.

For topical application to be used in the lower intestinal tract or vaginally, a rectal suppository, a suitable enema, a gel, an ointment, a solution, a suspension or an insert can be used. Topical transdermal patches may also be used. Transdermal patches have the added advantage of providing controlled delivery of the compositions of the invention to the body. Such dosage forms can be made by dissolving or dispersing the agent in the proper medium.

Compositions of the invention can be administered in the form of suppositories for rectal or vaginal administration. These can be prepared by mixing the agent with a suitable non-irritating carrier which is solid at room temperature but liquid at rectal temperature and therefore will melt in the rectum or vagina to release the drug. Such materials include cocoa butter, beeswax, polyethylene glycols, a suppository wax or a salicylate, and which is solid at room temperature, but liquid at body temperature and, therefore, will melt in the rectum or vaginal cavity and release the active agent.

Compositions which are suitable for vaginal administration also include pessaries, tampons, creams, gels, pastes, foams, films, or spray compositions containing such carriers as are known in the art to be appropriate. The carrier employed in the

sucrose octasulfate /contraceptive agent should be compatible with vaginal administration and/or coating of contraceptive devices. Combinations can be in solid, semi-solid and liquid dosage forms, such as diaphragm, jelly, douches, foams, films, ointments, creams, balms, gels, salves, pastes, slurries, vaginal suppositories, sexual lubricants, and coatings for devices, such as condoms, contraceptive sponges, cervical caps and diaphragms.

For ophthalmic applications, the pharmaceutical compositions can be formulated as micronized suspensions in isotonic, pH adjusted sterile saline, or, preferably, as solutions in isotonic, pH adjusted sterile saline, either with or without a preservative such as benzylalkonium chloride. Alternatively, for ophthalmic uses, the compositions can be formulated in an ointment such as petrolatum. Exemplary ophthalmic compositions include eye ointments, powders, solutions and the like.

Powders and sprays can contain, in addition to sucrose octasulfate and/or antibiotic or contraceptive agent(s), carriers such as lactose, talc, aluminum hydroxide, calcium silicates and polyamide powder, or mixtures of these substances. Sprays can additionally contain customary propellants, such as chlorofluorohydrocarbons and volatile unsubstituted hydrocarbons, such as butane and propane.

Ordinarily, an aqueous aerosol is made by formulating an aqueous solution or suspension of the agent together with conventional pharmaceutically acceptable carriers and stabilizers. The carriers and stabilizers vary with the requirements of the particular compound, but typically include nonionic surfactants (Tweens, Pluronics, or polyethylene glycol), proteins like serum albumin, sorbitan esters, oleic acid, lecithin, amino acids such as glycine, buffers, salts, sugars or sugar alcohols. Aerosols generally are prepared from isotonic solutions.

Compositions of the invention can also be orally administered in any orally-acceptable dosage form including, but not limited to, capsules, cachets, pills, tablets, lozenges (using a flavored basis, usually sucrose and acacia or tragacanth), powders, granules, or as a solution or a suspension in an aqueous or non-aqueous liquid, or as an oil-in-water or water-in-oil liquid emulsion, or as an elixir or syrup, or as pastilles (using an inert base, such as gelatin and glycerin, or sucrose and acacia) and/or as mouth washes and the like, each containing a predetermined amount of sucrose octasulfate and/or antibiotic or contraceptive agent(s) as an active ingredient. A compound may also be administered as a bolus, electuary or paste. In the case of tablets for oral use, carriers which are commonly used include lactose and corn starch. Lubricating agents, such as magnesium stearate, are also typically added. For oral administration in a capsule form, useful diluents include lactose and dried corn starch. When aqueous suspensions are required for oral use, the active ingredient is combined

with emulsifying and suspending agents. If desired, certain sweetening, flavoring or coloring agents may also be added.

Tablets, and other solid dosage forms, such as dragees, capsules, pills and granules, may be scored or prepared with coatings and shells, such as enteric coatings and other coatings well known in the pharmaceutical-formulating art. They may also be formulated so as to provide slow or controlled release of the active ingredient therein using, for example, hydroxypropylmethyl cellulose in varying proportions to provide the desired release profile, other polymer matrices, liposomes and/or microspheres. They may be sterilized by, for example, filtration through a bacteria-retaining filter, or by incorporating sterilizing agents in the form of sterile solid compositions which can be dissolved in sterile water, or some other sterile injectable medium immediately before use. These compositions may also optionally contain opacifying agents and may be of a composition that they release the active ingredient(s) only, or preferentially, in a certain portion of the gastrointestinal tract, optionally, in a delayed manner. Examples of embedding compositions which can be used include polymeric substances and waxes. The active ingredient can also be in micro-encapsulated form, if appropriate, with one or more of the above-described excipients.

Liquid dosage forms for oral administration include pharmaceutically acceptable emulsions, microemulsions, solutions, suspensions, syrups and elixirs. In addition to the active ingredient, the liquid dosage forms may contain inert diluents commonly used in the art, such as, for example, water or other solvents, solubilizing agents and emulsifiers, such as ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate, propylene glycol, 1,3-butylene glycol, oils (in particular, cottonseed, groundnut, corn, germ, olive, castor and sesame oils), glycerol, tetrahydrofuryl alcohol, polyethylene glycols and fatty acid esters of sorbitan, and mixtures thereof.

Besides inert diluents, the oral compositions can also include adjuvants such as wetting agents, emulsifying and suspending agents, sweetening, flavoring, coloring, perfuming and preservative agents.

Suspensions, in addition to the antiinfective agent(s) may contain suspending agents as, for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar and tragacanth, and mixtures thereof.

Sterile injectable forms of the compositions of this invention can be aqueous or oleaginous suspension. These suspensions may be formulated according to techniques known in the art using suitable dispersing or wetting agents and suspending agents. Wetting agents, emulsifiers and lubricants, such as sodium lauryl sulfate and

magnesium stearate, as well as coloring agents, release agents, coating agents, sweetening, flavoring and perfuming agents, preservatives and antioxidants can also be present in the compositions.

The sterile injectable preparation may also be a sterile injectable solution or suspension in a nontoxic parenterally-acceptable diluent or solvent, for example as a solution in 1,3-butanediol. Among the acceptable vehicles and solvents that may be employed are water, Ringer's solution and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose, any bland fixed oil may be employed including synthetic mono-or di-glycerides. Fatty acids, such as oleic acid and its glyceride derivatives are useful in the preparation of injectables, as are natural pharmaceutically-acceptable oils, such as olive oil or castor oil, especially in their polyoxyethylated versions. These oil solutions or suspensions may also contain a long-chain alcohol diluent or dispersant, such as Ph. Helv or similar alcohol.

The antiinfective agent or a pharmaceutically acceptable salt thereof will represent some percentage of the total dose in other dosage forms in a material forming a combination product, including liquid solutions or suspensions, suppositories, douches, enemas, gels, creams, emulsions, lotions slurries, soaps, shampoos, detergents, powders, sprays, lipsticks, foams, pastes, toothpastes, ointments, salves, balms, douches, drops, troches, lozenges, mouthwashes, rinses and others. Creams and gels for example, are typically limited by the physical chemical properties of the delivery medium to concentrations less than 20% (*e.g.*, 200 mg/gm). For special uses, far less concentrated preparations can be prepared, (*e.g.*, lower percent formulations for pediatric applications). For example, the pharmaceutical composition of the invention can comprise sucrose octasulfate in an amount of 0.001-99%, typically 0.01-75%, more typically 0.1-20%, especially 1-10% by weight of the total preparation. In particular, a preferred concentration thereof in the preparation is 0.5-50%, especially 0.5-25%, such as 1-10%. It can be suitably applied 1-10 times a day, depending on the type and severity of the condition to be treated or prevented.

Given the low toxicity of an antiinfective agent or a pharmaceutically acceptable salt thereof over many decades of clinical use as an anti-ulcerant [W.R. Garnett, *Clin. Pharm.* 1:307-314 (1982); R.N. Brogden et al., *Drugs* 27:194-209 (1984); D.M. McCarthy, *New Eng J Med.*, 325:1017-1025 (1991), an upper limit for the therapeutically effective dose is not a critical issue.

For prophylactic applications, the pharmaceutical composition of the invention can be applied prior to potential infection. The timing of application prior to potential infection can be optimized to maximize the prophylactic effectiveness of the

compound. The timing of application will vary depending on the mode of administration, the epithelial surface to which it is applied, the surface area, doses, the stability and effectiveness of composition under the pH of the epithelial surface, the frequency of application, *e.g.*, single application or multiple applications. One skilled in the art will be able to determine the most appropriate time interval required to maximize prophylactic effectiveness of the compound.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, genetics, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Genetics*; *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, J. *et al.* (Cold Spring Harbor Laboratory Press (1989)); *Short Protocols in Molecular Biology*, 3rd Ed., ed. by Ausubel, F. *et al.* (Wiley, NY (1995)); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed. (1984)); Mullis *et al.* U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. (1984)); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y); *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London (1987)); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds. (1986)); and Miller, J. *Experiments in Molecular Genetics* (Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1972)).

X. The Role of Transcription Activation Factor Polypeptides in Biofilms

In one embodiment, the invention pertains to a method for dispersing or preventing the formation of a biofilm on a surface or in an area, by administering an effective amount of a transcription factor modulating compound, e.g., a HTH protein modulating compound, an AraC family polypeptide modulating compound, a MarA family polypeptide modulating compound, or a MarA inhibiting compound.

It has been discovered that the absence of MarA and its homologs has a negative effect on biofilm formation in *E. coli*. In order to confirm this finding genetically, plasmid encoded *marA* was transformed into an *E. coli* strain deleted of *marA*, *soxS*, and *rob* (triple knockout). The expression of MarA in this triple knockout restored biofilm formation in this host to a level that was comparable to that of the wild type host.

The term "biofilm" includes biological films that develop and persist at interfaces in aqueous and other environments. Biofilms are composed of microorganisms embedded in an organic gelatinous structure composed of one or more matrix polymers which are secreted by the resident microorganisms. The term "biofilm"

also includes bacteria that are attached to a surface in sufficient numbers to be detected or communities of microorganisms attached to a surface (Costerton, J. W., *et al.* (1987) *Ann. Rev. Microbiol.* 41:435-464; Shapiro, J. A. (1988) *Sci Am.* 256:82-89; O'Toole, G. *et al.* (2000) *Annu Rev Microbiol.* 54:49-79).

In another embodiment, the invention pertains to methods of treating biofilm associated states in a subject, by administering to said subject an effective amount of a transcription factor modulating compound, e.g., a MarA family inhibiting compound, such that the biofilm associated state is treated.

The term "biofilm associated states" includes disorders which are characterized by the presence or potential presence of a bacterial biofilm. Many medically important pathogens form biofilms and biofilm formation is often one component of the infectious process (Costerton, J. W. *et al.* (1999) *Science* 284:1318-1322). Examples of biofilm associated states include, but are not limited to, middle ear infections, cystic fibrosis, osteomyelitis, acne, dental cavities, and prostatitis. Biofilm associated states also include infection of the subject by one or more bacteria, *e.g.*, *Pseudomonas aeruginosa.* One consequence of biofilm formation is that bacteria within biofilms are generally less susceptible to a range of different antibiotics relative to their planktonic counterparts.

Furthermore, the invention also pertains to methods for preventing the formation of biofilms on surfaces or in areas, by contacting the area with an effective amount of a transcription factor modulating compound, e.g., a MarA family inhibiting compound, etc.

Industrial facilities employ many methods of preventing biofouling of industrial water systems. Many microbial organisms are involved in biofilm formation in industrial waters. Growth of slime-producing bacteria in industrial water systems causes problems including decreased heat transfer, fouling and blockage of lines and valves, and corrosion or degradation of surfaces. Control of bacterial growth in the past has been accomplished with biocides. Many biocides and biocide formulations are known in the art. However, many of these contain components which may be environmentally deleterious or toxic, and are often resistant to breakdown.

The transcription factor inhibiting compounds, such as but not limited to AraC family inhibiting compounds and MarA family inhibiting compounds, of the present invention are useful in a variety of environments including industrial, clinical, the household, and personal care. The compositions of the invention may comprise one or more compounds of the invention as an active ingredient acting alone, additively, or synergistically against the target organism.

The MarA family inhibiting compounds and modulating compounds of the invention may be formulated in a composition suitable for use in environments including industry, pharmaceutics, household, and personal care. In an embodiment, the compounds of the invention are soluble in water. The modulating compounds may be applied or delivered with an acceptable carrier system. The composition may be applied or delivered with a suitable carrier system such that the active ingredient (e.g., transcription factor modulating compound of the invention such as a MarA family modulating compound, e.g., a MarA family polypeptide inhibiting compound) may be dispersed or dissolved in a stable manner so that the active ingredient, when it is administered directly or indirectly, is present in a form in which it is available in a advantageous way.

Also, the separate components of the compositions of the invention may be preblended or each component may be added separately to the same environment according to a predetermined dosage for the purpose of achieving the desired concentration level of the treatment components and so long as the components eventually come into intimate admixture with each other. Further, the present invention may be administered or delivered on a continuous or intermittent basis.

A transcription factor modulating compound, e.g., a MarA family modulating compound of the present invention, when present in a composition will generally be present in an amount from about 0.000001% to about 100%, more preferably from about 0.001% to about 50%, and most preferably from about 0.01% to about 25%.

For compositions of the present invention comprising a carrier, the composition comprises, for example, from about 1% to about 99%, preferably from about 50% to about 99%, and most preferably from about 75% to about 99% by weight of at least one carrier.

The transcription factor modulating compound, e.g., the MarA family polypeptide inhibiting compound, of the invention may be formulated with any suitable carrier and prepared for delivery in forms, such as, solutions, microemulsions, suspensions or aerosols. Generation of the aerosol or any other means of delivery of the present invention may be accomplished by any of the methods known in the art. For example, in the case of aerosol delivery, the compound is supplied in a finely divided form along with any suitable carrier with a propellant. Liquefied propellants are typically gases at ambient conditions and are condensed under pressure. The propellant may be any acceptable and known in the art including propane and butane, or other lower alkanes, such as those of up to 5 carbons. The composition is held within a

container with an appropriate propellant and valve, and maintained at elevated pressure until released by action of the valve.

The compositions of the invention may be prepared in a conventional form suitable for, but not limited to topical or local application such as an ointment, paste, gel, spray and liquid, by including stabilizers, penetrants and the carrier or diluent with the compound according to a known technique in the art. These preparations may be prepared in a conventional form suitable for enteral, parenteral, topical or inhalational applications.

The present invention may be used in compositions suitable for household use. For example, compounds of the present invention are also useful as active antimicrobial ingredients in household products such as cleansers, detergents, disinfectants, dishwashing liquids, soaps and detergents. In an embodiment, the transcription factor modulating compound of the present invention may be delivered in an amount and form effective for the prevention, removal or termination of microbes.

The compositions of the invention for household use comprise, for example, at least one transcription factor modulating compound of the invention and at least one suitable carrier. For example, the composition may comprise from about 0.00001% to about 50%, preferably from about 0.0001% to about 25%, most preferably from about 0.0005% to about 10% by weight of the modulating compound based on the weight percentage of the total composition.

The transcription factor modulating compound of the present invention may also be used in hygiene compositions for personal care. For instance, compounds of the invention can be used as an active ingredient in personal care products such as facial cleansers, astringents, body wash, shampoos, conditioners, cosmetics and other hygiene products. The hygiene composition may comprise any carrier or vehicle known in the art to obtain the desired form (such as solid, liquid, semisolid or aerosol) as long as the effects of the compound of the present invention are not impaired. Methods of preparation of hygiene compositions are not described herein in detail, but are known in the art. For its discussion of such methods, The CTFA Cosmetic Ingredient Handbook, Second Edition, 1992, and pages 5-484 of A Formulary of Cosmetic Preparations (Vol. 2, Chapters 7-16) are incorporated herein by reference.

The hygiene composition for use in personal care comprise generally at least one modulating compound of the present application and at least one suitable carrier. For example, the composition may comprise from about 0.00001% to about 50%, preferably from about 0.0001% to about 25%, more preferably from about 0.0005% to about 10% by weight of the transcription factor modulating compound of the invention based on the weight percentage of the total composition.

The transcription factor modulating compound of the present invention may be used in industry. In the industrial setting, the presence of microbes can be problematic, as microbes are often responsible for industrial contamination and biofouling. Compositions of the invention for industrial applications may comprise an effective amount of the compound of the present invention in a composition for industrial use with at least one acceptable carrier or vehicle known in the art to be useful in the treatment of such systems. Such carriers or vehicles may include diluents, deflocculating agents, penetrants, spreading agents, surfactants, suspending agents, wetting agents, stabilizing agents, compatibility agents, sticking agents, waxes, oils, co-solvents, coupling agents, foams, antifoaming agents, natural or synthetic polymers, elastomers and synergists. Methods of preparation, delivery systems and carriers for such compositions are not described here in detail, but are known in the art. For its discussion of such methods, U.S. Patent No. 5,939,086 is herein incorporated by reference. Furthermore, the preferred amount of the composition to be used may vary according to the active ingredient(s) and situation in which the composition is being applied.

The transcription factor modulating compounds, *e.g.*, MarA family polypeptide inhibiting compounds, and compositions of the present invention may be useful in nonaqueous environments. Such nonaqueous environments may include, but are not limited to, terrestrial environments, dry surfaces or semi-dry surfaces in which the compound or composition is applied in a manner and amount suitable for the situation.

The transcription factor modulating compounds, *e.g.*, MarA family polypeptide modulating compounds, *e.g.*, MarA inhibiting compounds, of the present invention may be used to form contact-killing coatings or layers on a variety of substrates including personal care products (such as toothbrushes, contact lens cases and dental equipment), healthcare products, household products, food preparation surfaces and packaging, and laboratory and scientific equipment. Further, other substrates include medical devices such as catheters, urological devices, blood collection and transfer devices, tracheotomy devices, intraocular lenses, wound dressings, sutures, surgical staples, membranes, shunts, gloves, tissue patches, prosthetic devices (e.g., heart valves) and wound drainage tubes. Still further, other substrates include textile products such as carpets and fabrics, paints and joint cement. A further use is as an antimicrobial soil fumigant.

The transcription factor modulating compounds of the invention may also be incorporated into polymers, such as polysaccharides (cellulose, cellulose derivatives, starch, pectins, alginate, chitin, guar, carrageenan), glycol polymers, polyesters,

polyurethanes, polyacrylates, polyacrylonitrile, polyamides (e.g., nylons), polyolefins, polystyrenes, vinyl polymers, polypropylene, silks or biopolymers. The modulating compounds may be conjugated to any polymeric material such as those with the following specified functionality: 1) carboxy acid, 2) amino group, 3) hydroxyl group and/or 4) haloalkyl group.

The composition for treatment of nonaqueous environments may be comprise at least one transcription factor modulating compound of the present application and at least one suitable carrier. In an embodiment, the composition comprises from about 0.001% to about 75%, advantageously from about 0.01% to about 50%, and preferably from about 0.1% to about 25% by weight of a transcription factor modulating compound of the invention based on the weight percentage of the total composition.

The transcription factor modulating compounds and compositions of the invention may also be useful in aqueous environments. "Aqueous environments" include any type of system containing water, including, but not limited to, natural bodies of water such as lakes or ponds; artificial, recreational bodies of water such as swimming pools and hot tubs; and drinking reservoirs such as wells. The compositions of the present invention may be useful in treating microbial growth in these aqueous environments and may be applied, for example, at or near the surface of water.

The compositions of the invention for treatment of aqueous environments may comprise at least one transcription factor modulating compound of the present invention and at least one suitable carrier. In an embodiment, the composition comprises from about 0.001% to about 50%, advantageously from about 0.003% to about 15%, preferably from about 0.01% to about 5% by weight of the compound of the invention based on the weight percentage of the total composition.

The present invention also provides a process for the production of an antibiofouling composition for industrial use. Such process comprises bringing at least one of any industrially acceptable carrier known in the art into intimate admixture with a transcription factor modulating compound of the present invention. The carrier may be any suitable carrier discussed above or known in the art.

The suitable antibiofouling compositions may be in any acceptable form for delivery of the composition to a site potentially having, or having at least one living microbe. The antibiofouling compositions may be delivered with at least one suitably selected carrier as hereinbefore discussed using standard formulations. The mode of delivery may be such as to have a binding inhibiting effective amount of the antibiofouling composition at a site potentially having, or having at least one living microbe. The antibiofouling compositions of the present invention are useful in treating

microbial growth that contributes to biofouling, such as scum or slime formation, in these aqueous environments. Examples of industrial processes in which these compounds might be effective include cooling water systems, reverse osmosis membranes, pulp and paper systems, air washer systems and the food processing industry. The antibiofouling composition may be delivered in an amount and form effective for the prevention, removal or termination of microbes.

The antibiofouling composition of the present invention generally comprise at least one compound of the invention. The composition may comprise from about 0.001% to about 50%, more preferably from about 0.003% to about 15%, most preferably from about 0.01% to about 5% by weight of the compound of the invention based on the weight percentage of the total composition.

The amount of antibiofouling composition may be delivered in an amount of about 1 mg/l to about 1000 mg/l, advantageously from about 2 mg/l to about 500 mg/l, and preferably from about 20 mg/l to about 140 mg/l.

Antibiofouling compositions for water treatment generally comprise transcription factor modulating compounds of the invention in amounts from about 0.001% to about 50% by weight of the total composition. Other components in the antibiofouling compositions (used at 0.1% to 50%) may include, for example, 2-bromo-2-nitropropane-1,3-diol (BNPD), β-nitrostyrene (BNS), dodecylguanidine hydrochloride, 2,2-dibromo-3-nitrilopropionamide (DBNPA), glutaraldehyde, isothiazolin, methylene bis(thiocyanate), triazines, n-alkyl dimethylbenzylammonium chloride, trisodium phosphate-based, antimicrobials, tributyltin oxide, oxazolidines, tetrakis (hydroxymethyl)phosphonium sulfate (THPS), phenols, chromated copper arsenate, zinc or copper pyrithione, carbamates, sodium or calcium hypochlorite, sodium bromide, halohydantoins (Br, Cl), or mixtures thereof.

Other possible components in the compositions of the invention include biodispersants (about 0.1% to about 15% by weight of the total composition), water, glycols (about 20-30%) or Pluronic (at approximately 7% by weight of the total composition). The concentration of antibiofouling composition for continuous or semi-continuous use is about 5 to about 70 mg/l.

Antibiofouling compositions for industrial water treatment may comprise compounds of the invention in amounts from about 0.001% to about 50% based on the weight of the total composition. The amount of compound of the invention in antibiofouling compositions for aqueous water treatment may be adjusted depending on the particular environment. Shock dose ranges are generally about 20 to about 140 mg/l; the concentration for semi-continuous use is about 0.5X of these concentrations.

The invention also pertains, at least in part, to a method of regulating biofilm development. The method includes administering a composition which contains a transcription factor modulating compound of the invention. The composition can also include other components which enhance the ability of the composition to degrade biofilms.

The composition can be formulated as a cleaning product, *e.g.*, a household or an industrial cleaner to remove, prevent, inhibit, or modulate biofilm development. Advantageously, the biofilm is adversely affected by the administration of the compound of the invention, *e.g.*, biofilm development is diminished. These compositions may include compounds such as disinfectants, soaps, detergents, as well as other surfactants. Examples of surfactants include, for example, sodium dodecyl sulfate; quaternary ammonium compounds; alkyl pyridinium iodides; TWEEN 80, TWEEN 85, TRITON X-100; BRIJ 56; biological surfactants; rhamnolipid, surfactin, visconsin, and sulfonates. The composition of the invention may be applied in known areas and surfaces where disinfection is required, including but not limited to drains, shower curtains, grout, toilets and flooring. A particular application is on hospital surfaces and medical instruments. The disinfectant of the invention may be useful as a disinfectant for bacteria such as, but not limited to, *Pseudomonadaceae*, *Azatobacteraceae*, *Rhizabiaceae*, *Mthylococcaceae*, *Halobacteriaceae*, *Acetobacteraceae*, *Legionellaceae*, *Neisseriaceae*, and other genera.

The invention also pertains to a method for cleaning and disinfecting contact lenses. The method includes contacting the contact lenses with a solution of at least one compound of the invention in an acceptable carrier. The invention also pertains to the solution comprising the compound, packaged with directions for using the solution to clean contact lenses.

The invention also includes a method of treating medical indwelling devices. The method includes contacting at least one compound of the invention with a medical indwelling device, such as to prevent or substantially inhibit the formation of a biofilm. Examples of medical indwelling devices include catheters, orthopedic devices and implants.

A dentifrice or mouthwash containing the compounds of the invention may be formulated by adding the compounds of the invention to dentifrice and mouthwash formulations, *e.g.*, as set forth in *Remington's Pharmaceutical Sciences*, 18th Ed., Mack Publishing Co., 1990, Chapter 109 (incorporated herein by reference in its entirety). The dentifrice may be formulated as a gel, paste, powder or slurry. The dentifrice may include binders, abrasives, flavoring agents, foaming agents and

humectants. Mouthwash formulations are known in the art, and the compounds of the invention may be advantageously added to them.

In one embodiment, the invention pertains to each of the transcription factor modulating compounds described herein in Tables 4 and 5, and in Formulae (I)-(X).

The contents of all references, patent applications and patents, cited throughout this application are hereby expressly incorporated by reference. Each reference disclosed herein is incorporated by reference herein in its entirety. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety.

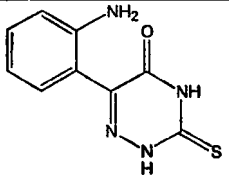
The invention is further illustrated by the following examples, which should not be construed as further limiting.

EXEMPLIFICATION OF THE INVENTION

Example 1: Synthesis of Test Compounds

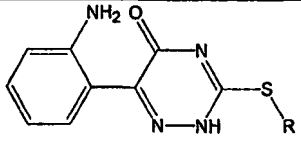
The transcriptional modulating compounds described in this application can be synthesized by art recognized techniques or using the methods described herein.

6-(2-Amino-phenyl)-3-thioxo-3,4-dihydro-2H-[1,2,4]triazin-5-one



This was prepared by a modified literature procedure (Doleschall, G.; Lempert, K. *Tetrahedron* 1973, *29*, 639-649). Isatin (10g, 67.96 mmol) was dissolved in ca. 10% aqueous KOH (9.9 g in 100 mL of water) and then treated with thiosemicarbazide (6.28 g; 68.90 mmol). After 1 hour of heating at 115 °C (bath temperature), the reaction mixture was poured over ice and treated with glacial acetic acid drop-wise, till the *p*H was ca. 5. The yellow fluffy precipitate was filtered, washed copiously with water (8 x 50 mL) and dried first in air and then under high vacuum to afford 12.9 g of yellow solid.

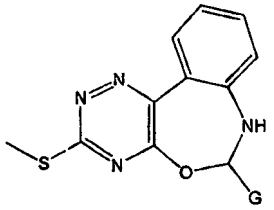
6-(2-Amino-phenyl)-3-butylsulfanyl-2H-[1,2,4]triazin-5-one



This was prepared by a modified literature procedure (Doleschall, G.; Lempert, K. *Tetrahedron* 1973, *29*, 639-649). The product from the previous experiment (8.0 g, 36.3 mmol) was dissolved in ca. 10% aq. KOH (10.3 g in 100 mL of water) and treated with ^{n}BuI (7 mL). Ethanol (70 mL) was added to it and the reaction mixture was allowed to stir overnight. The reaction mixture was diluted with ether (100 mL) and water (70 mL). The ether layer was separated and the aqueous layer washed further with ether (3 x 100 mL) and then poured over ice. Upon careful, drop-wise addition of glacial acetic acid with vigorous stirring at 0-4 °C, yellow precipitate was obtained which was filtered, washed with water (4 x 20 mL) and then with ether (2 x 10mL) and dried. Yield: 5.12 g.

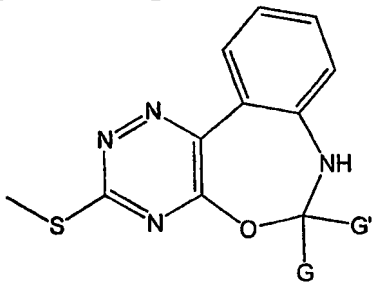
Other alkyl or substituted alkyl halides were used instead of n-butyliodide following the similar method.

3-Methylsulfanyl-6-(G)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene



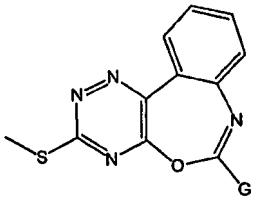
This was prepared by a modified literature procedure (Doleschall, G.; Lempert, K. *Tetrahedron* 1973, *29*, 639-649). To a suspension of compound **2** (or analogs of **2**) (ca. 0.384 mmol, 1 equiv) in dry ethanol (3-4 mL), 25 µL of glacial acetic acid was added, followed by ca. 1.1 equiv of the corresponding aldehyde (**G**-CHO, where, **G** = substituted or unsubstituted aliphatic, aromatic, or heterocyclic groups). The reaction mixture was refluxed for ca. 5-7 min resulting in a dark red – dark-reddish orange solution. Upon cooling to room temperature orange-orange-yellow solid crashed out of solution, which was filtered, washed with cold (ca. –30 °C) methanol (2 x 1 mL), and/or ether and dried. In some cases, the crude products were recrystallized from DMF/ether or methanol/ether; in most of the cases, the crude products, prepared as above, were >95% pure. Various ketones (GCOG') were reacted with **2** (or analogs of **2**) in a similar way to afford compounds of structural type **4**. All the final compounds were characterized by means of ^{1}H NMR, LC-MS, HPLC (C_{18} columns, acetonitrile/water with 0.01% triethylamine as mobile phase), and CHN analyses.

General synthesis of orthoesters, $GC(OR)_3$; R = Me



The syntheses of the desired orthoesters were accomplished by a modified literature procedure in multiple steps (McClelland, R. A. *et al. J. Org. Chem.* 1981, *46*, 1011-1012). Several novel orthoesters were prepared by this method. To a solution of an acid chloride in dichloromethane, *N*-methylaniline was added slowly, followed by triethylamine and catalytic amount of 4-dimethylaminopyridine. After stirring it for ca. 12 h, the reaction mixture was diluted with ether, the precipitate was filtered, washed with ether and dried. The amide, thus prepared, was then stirred overnight with methyl triflate in dichloromethane, diluted with ether, and the precipitate was filtered, washed, and dried to obtain an imidatonium triflate salt. This salt was dissolved in dichloromethane, cooled to 0 °C, and added slowly, with stirring, to a cold (0 °C) solution of sodium methoxide in dry methanol over a period of ca. 30-60 min. The solvent was evaporated to dryness and the residue was extracted in n-hexane. Upon evaporation of hexane, the white solid was obtained, which was dissolved in dry methanol and treated with glacial acetic acid. After 10 minutes of stirring, the excess acid was neutralized with potassium carbonate (solid), and the solvent removed under vacuum. The residue was extracted in ether, washed with water, and dried over potassium carbonate. The crude material was obtained by evaporation of ether, and further purified either by flash chromatography or fractional distillation.

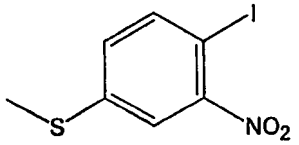
3-Methylsulfanyl-6-(G)-5-oxa-1,2,4,7-tetraaza-dibenzo[a,c]cycloheptene



Compound of the type **2** (0.384 mmol, 1 equiv) was suspended in ethanol (2-3 mL), treated with glacial acetic acid (100 μL), followed by an orthoester (2 equiv) of the general formula G-$C(OR)_3$, where G = substituted or unsubstituted aliphatic, aromatic, or heterocyclic group, R = H, substituted or unsubstituted aliphatic, aromatic, or heterocyclic group. The reaction mixture was refluxed for 70-180 minutes, cooled to room temperature. In some cases, the product crashed out of solution, in others, the

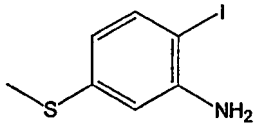
crude reaction mixture was evaporated to dryness, re-dissolved in a minimum amount of methanol and diluted with ether. The solid was washed with ether (cold, 0-4 °C; 1x 1mL) and dried under vacuum.

4-Iodo-3-nitrothioanisole



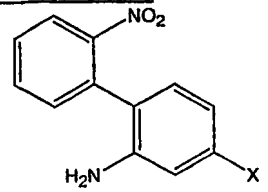
A flask was charged with 10 g of 4-iodothioanisole and 5.5 mL of dimethylsulfate and warmed to ca. 90 °C for 10 min. The resulting solution was dissolved in conc. sulfuric acid (30 mL) and cooled to ca. 0-4 °C, whereupon it was treated, with extreme caution, slowly with conc. HNO_3 (ca. 2 mL) while maintaining the reaction temperature below 4 °C. After stirring it for ca. 10 min, the reaction mixture was stirred at ca. 90 °C for ca. 3 d. The reaction was monitored with HPLC, TLC, and LC-MS, and if needed, smaller portions of nitric acid were added to the reaction mixture to force it to completion. Use of fuming sulfuric acid is also helpful. After the complete consumption of the aromatic starting material, the reaction mixture was cooled, poured over crushed ice, treated with 30% aq. perchloric acid at 4 °C. The light colored precipitate was filtered, washed thoroughly with water, and dried under vacuum. The perchlorate salt was stirred with saturated aq. NaCl solution at 95 °C for 3-6 h. Upon cooling to room temperature, the precipitate was filtered, washed thoroughly with water to get rid of any inorganic salts, and dried under vacuum to obtain 4-iodo-3-nitrothioanisole in > 80% yield.

4-Iodo-3-aminothioanisole



Ca. 7 g of 4-iodo-3-nitrothioanisole was dissolved in absolute ethanol and treated slowly with a solution of $SnCl_2.2H_2O$ in 12% aq. HCl. The reaction mixture was stirred at 50 °C for 25-30 min., when the HPLC monitoring indicated that the reaction was complete. The reaction mixture was poured over crushed ice, and treated with aq. NaOH solution to *p*H 8. The precipitate was filtered, washed with water and dried in air. The crude material was crystallized by overnight cooling (4 °C) of its ethanol (minimum amount) solution treated with 10% aq. HCl. The crystalline material was further dried under high vacuum to afford the desired amine as its hydrochloride salt in >70% yield.

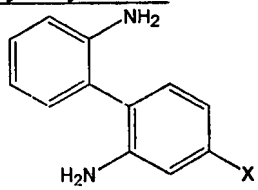
4-Methylsulfanyl-2'-nitro-biphenyl-2-ylamine



X = H, -OR, -SR

A methanol/dioxane (20 mL/5 mL of methanol/dioxane) solution/suspension of 4-iodo-3-nitrothioanisole (1 mmol), $Pd(OAc)_2$ (0.01 mmol) was purged with argon for ~5 min. To this solution Et_3N (3 mmol), and 5 mL of water were added and purged with argon for another 5 min. To the above solution, 2-aminophenyl boronic acid (2 mmol, solution in 5 mL of DMF, purged with argon), was added and the reaction mixture was heated at 70 °C (oil bath temperature) for 2 h. The reaction was monitored by HPLC and LC-MS to follow the product formation. The reaction mixture was then cooled down to room temperature and filtered through diatomaceous earth. The filtrate was concentrated and purified using preparative HPLC.

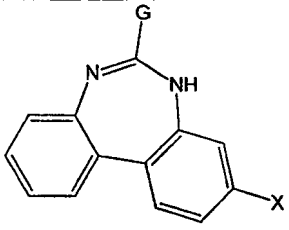
4-Methylsulfanyl-2'-amino-biphenyl-2-ylamine



X = H, -OR, -SR

A flask was charged with ca. 1 mmol of 4-Methylsulfanyl-2'-nitro-biphenyl-2-ylamine, 15 mL of ethanol, and 0.1 mmol of PtO_2, and stirred under hydrogen atmosphere at 40 psi for 10 minutes. The reaction mixture was filtered through diatomaceous earth, washed with ethanol, and the combined organic layer was evaporated to dryness. The crude material was purified by preparative HPLC. The same material can also be prepared by the previous method, (Suzuki coupling conditions) starting from 4-iodo-3-aminothioanisole, and purified by preparative HPLC.

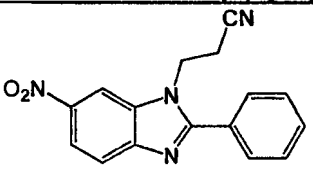
6-(G)-3-methylsulfanyl-5H-dibenzo[d,f][1,3]diazepine

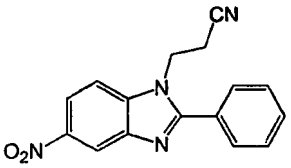


X = H, -OR, -SR

To a solution of the 2,2'-biphenyldiamine (0.093 g; 0.51 mmol) in ethanol (2 mL), were added glacial acetic acid (50 μL) and 2 equiv of an orthoester of the general formula $GC(OR)_3$. In case of TFA salt of the diamine, there was no need of adding acetic acid to the reaction mixture. The reaction mixture was refluxed for 3h, cooled to room temperature, and evaporated to dryness. The residue was suspended in methanol saturated with dry HCl, stirred for a few minutes, filtered, washed with methanol, and finally with ether. The hydrochloride salt of the diazepine was dried under vacuum to afford a light yellow solid. In order to obtain a free base of the diazepine, the above hydrochloride salt was suspended in methanol, and treated with 10% aq. NaOH solution. After stirring at room temperature for ca. 10 min, the precipitate was filtered, washed with water, and dried under vacuum.

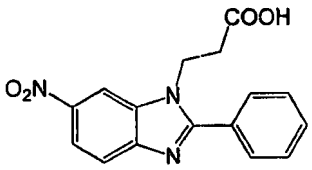
3-(6-Nitro-2-phenyl-benzoimidazol-1-yl)-propionitrile and 3-(5-Nitro-2-phenyl-benzoimidazol-1-yl)-propionitrile

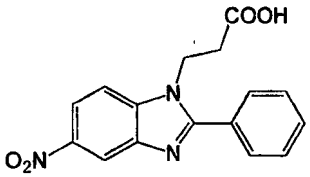




A mixture of 5-nitro-2-phenylbenzimidazole (1g, 4.2 mmol), acrylonitrile (50 mL) and N, N-dimethylpiridine (25 mg) was heated at 70 °C for 4 hr. The excess of acrylonitrile was evaporated and oily residue was subjected to the flash chromatography on silica gel using hexane-ethyl acetate (75: 25 v/v) as an eluent. Structure of the regioisomers was determined using ^1H NOESY studies. 0.25 g (20%) of the 6-nitro isomer and 0.23 g (18.9 %) of the 5-nitro isomer were obtained.

3-(6-Nitro-2-phenyl-benzoimidazol-1-yl)-propionic acid and 3-(5-Nitro-2-phenyl-benzoimidazol-1-yl)-propionic acid.





To the 6-nitro nitrile (0.15 g, 0.51 mmol) concentrated HCl (5 mL) was added and resulting mixture was heated at 50 °C for 0.5 hours. The acid was evaporated *in vacuo* and product was purified by HPLC. Yield 34 mg (21 %). An identical procedure was used starting from the 5-nitro nitrile yielding the product (22 mg, 13.8 %).

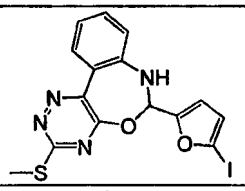
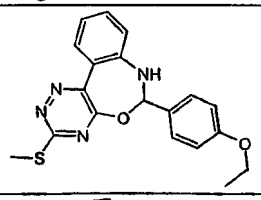
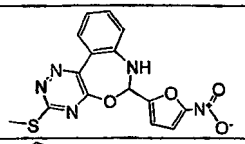
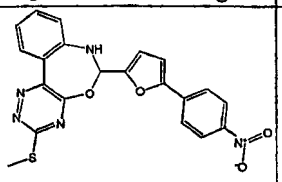
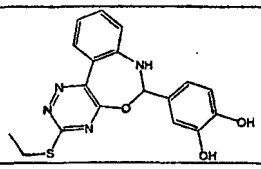
Example 2:

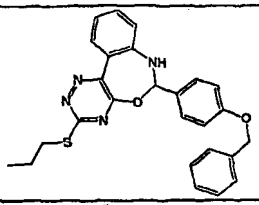
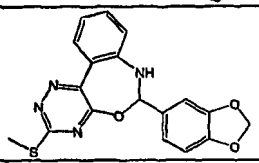
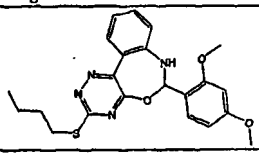
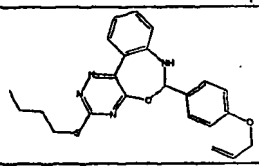
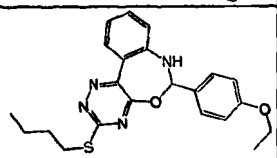
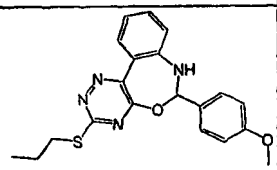
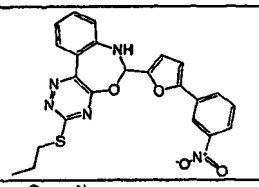
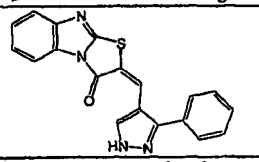
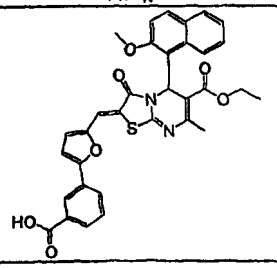
In this example, the expression of a selective marker (*e.g.*, ccdB) is put under the direct control of a promoter activated by MarA (*e.g.*, inaA, galT, or micF). In the absence of MarA, the expression of the selective marker is silent and cells survive. Synthesis of MarA from an inducible plasmid in a bacterial or yeast cell leads to the activation of the inaA promoter, expression of ccdB, and subsequently results in cell death. Compounds that inhibit MarA are those that permit cell survival under conditions of MarA expression. The results of this assay are given in Table 4. In Table 4, * indicates good inhibition of MarA and ** indicates very good inhibition of MarA.

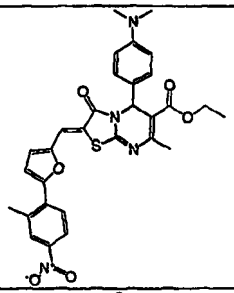
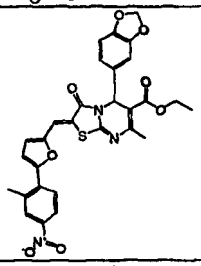
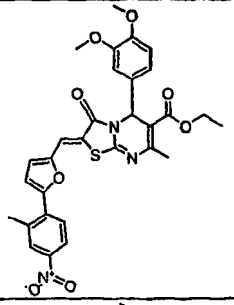
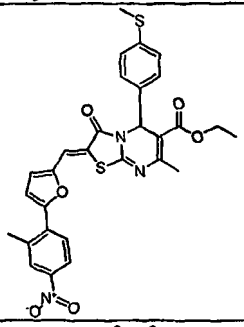
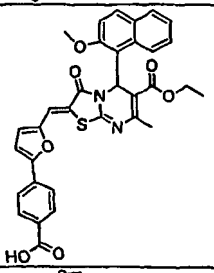
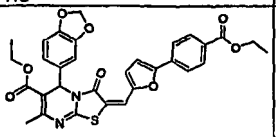
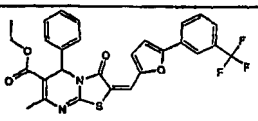
Example 3:

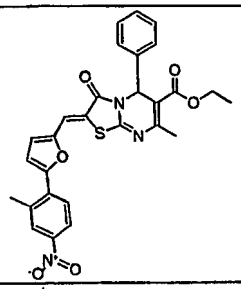
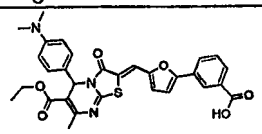
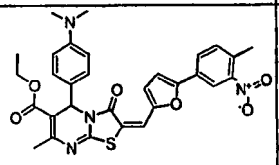
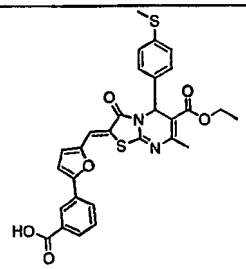
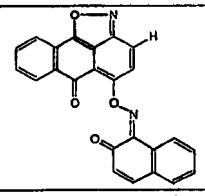
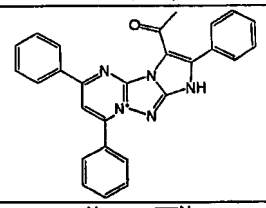
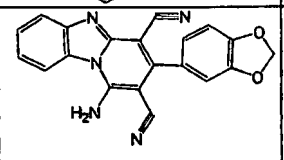
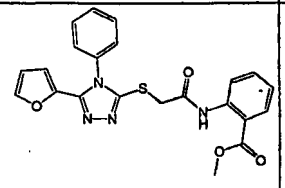
In this example, the expression of luciferase is put under the direct control of a promoter activated by MarA (*e.g.*, inaA, galT, or micF) in a cell constitutively expressing MarA. In the absence of MarA, cells luminesce. Upon modulating of MarA activity, the expression of the reporter is altered.

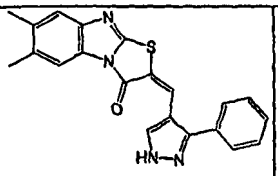
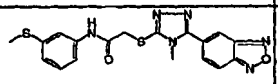
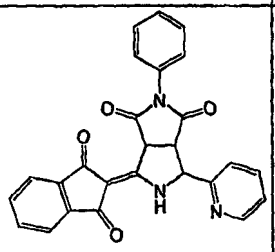
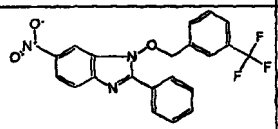
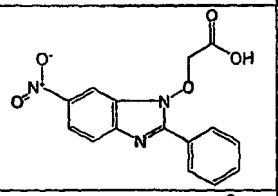
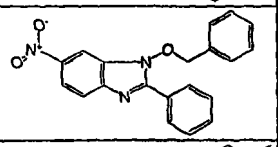
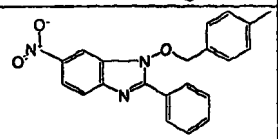
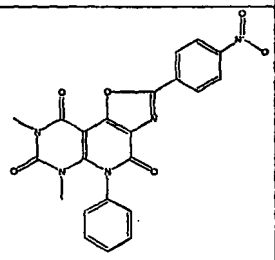
TABLE 4

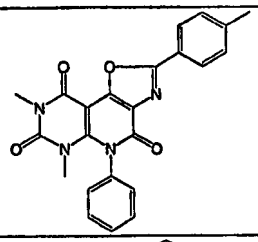
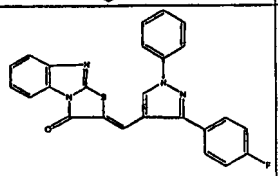
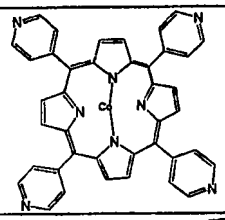
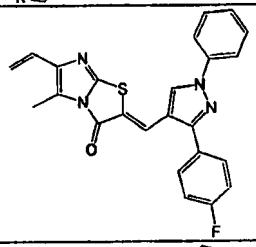
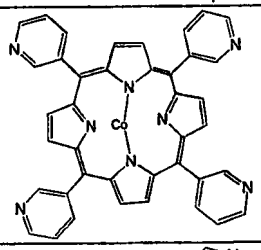
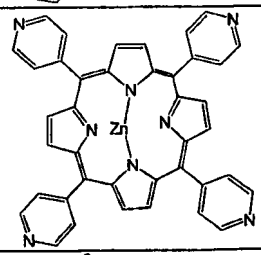
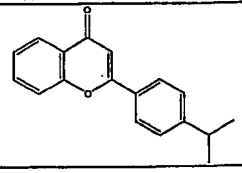
ID	Structure	Name	Affinity
A		6-(5-Iodo-furan-2-yl)-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene	**
B		6-(4-Ethoxy-phenyl)-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene	*
C		3-Methylsulfanyl-6-(5-nitro-furan-2-yl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene	*
D		3-Methylsulfanyl-6-[5-(4-nitro-phenyl)-furan-2-yl]-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo [a,c] cycloheptene	*
E		4-(3-Ethylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo [a,c] cyclohepten-6-yl)-benzene-1,2-diol	*

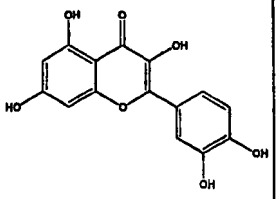
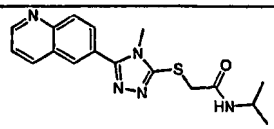
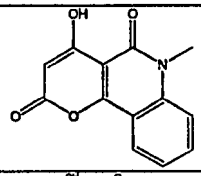
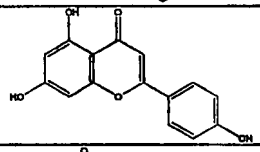
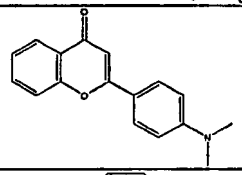
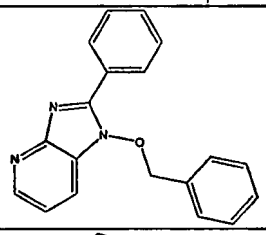
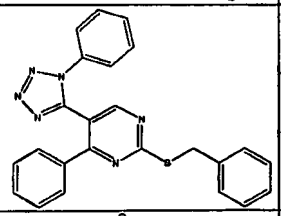
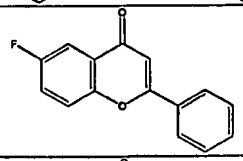
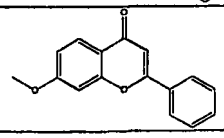
F		6-(4-Benzyloxy-phenyl)-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene	**
G		6-Benzo[1,3]dioxol-5-yl-3-methylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene	???
H		3-Butylsulfanyl-6-(2,4-dimethoxy-phenyl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo[a,c] cycloheptene	*
I		6-(4-Allyloxy-phenyl)-3-butylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo[a,c]cycloheptene	**
J		3-Butylsulfanyl-6-(4-ethoxy-phenyl)-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c] cycloheptene	*
K		6-(4-Methoxy-phenyl)-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza-dibenzo [a,c]cycloheptene	*
L		6-[5-(3-Nitro-phenyl)-furan-2-yl]-3-propylsulfanyl-6,7-dihydro-5-oxa-1,2,4,7-tetraaza -dibenzo[a,c] cycloheptene	**
M		2-(3-Phenyl-1H-pyrazol-4-ylmethylene)-benzo[4,5] imidazo[2,1-b]thiazol-3-one	**
N		2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-5-(2-methoxy-naphthalen-1-yl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	**

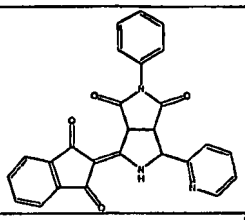
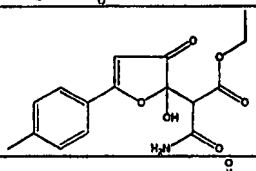
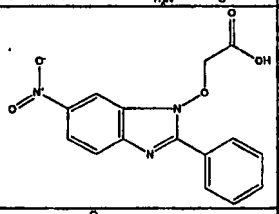
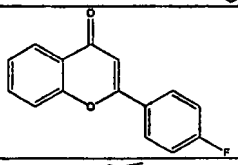
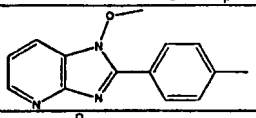
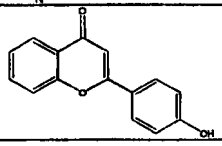
O		5-(4-Dimethylamino-phenyl)-7-methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro -5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester	***
P		5-Benzo[1,3]dioxol-5-yl-7-methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro -5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester	*
Q		5-(3,4-Dimethoxy-phenyl)-7-methyl-2-[5-(2-methyl-4-nitro -phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	***
R		7-Methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-ylmethylene]-5-(4-methyl sulfanyl-phenyl)-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a] pyrimidine-6-carboxylic acid ethyl ester	***
S		2-[5-(4-Carboxy-phenyl)-furan-2-ylmethylene]-5-(2-methoxy-naphthalen-1-yl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	***
T		5-Benzo[1,3]dioxol-5-yl-2-[5-(4-ethoxycarbonyl-phenyl)-furan-2-ylmethylene]-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	***
U		7-Methyl-3-oxo-5-phenyl-2-[5-(3-trifluoromethyl-phenyl)-furan-2-ylmethylene]-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	**

V		7-Methyl-2-[5-(2-methyl-4-nitro-phenyl)-furan-2-yl methylene]-3-oxo-5-phenyl-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	***
W		2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-5-(4-dimethylamino-phenyl)-7-methyl-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	**
X		5-(4-Dimethylamino-phenyl)-7-methyl-2-[5-(4-methyl-3-nitro-phenyl)-furan-2-yl methylene]-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	**
Y		2-[5-(3-Carboxy-phenyl)-furan-2-ylmethylene]-7-methyl-5-(4-methylsulfanyl-phenyl)-3-oxo-2,3-dihydro-5H-thiazolo[3,2-a]pyrimidine-6-carboxylic acid ethyl ester	***
Z		[1,2]Naphthoquinone 1-[O-(6-oxo-6H-anthra[1,9-cd] isoxazol-5-yl)-oxime]	*
AA		3-Acetyl-2,5,7-triphenyl-1H-1,3a,4,8-tetraaza-7a-azonia-cyclopenta[a]indene	***
AB		1-Amino-3-benzo[1,3]dioxol-5-yl-benzo[4,5]imidazo[1,2-a] pyridine-2,4-dicarbonitrile	**
AC		2-[2-(5-Furan-2-yl-4-phenyl-4H-[1,2,4]triazol-3-yl sulfanyl)-acetylamino]-benzoic acid methyl ester	*

AD		6,7-Dimethyl-2-(3-phenyl-1H-pyrazol-4-ylmethylene)-benzo[4,5]imidazo[2,1-b] thiazol-3-one	*
AE		2-(5-Benzo[1,2,5]oxadiazol-5-yl-4-methyl-4H-[1,2,4] triazol-3-ylsulfanyl)-N-(3-methylsulfanyl-phenyl)-acetamide	*
AF		4-(1,3-Dioxo-indan-2-ylidene)-2-phenyl-6-pyridin-2-yl-tetrahydro-pyrrolo[3,4-c] pyrrole-1,3-dione	**
AG		6-Nitro-2-phenyl-1-(3-trifluoromethyl-benzyloxy)-1H-benzoimidazole	**
AH		(6-Nitro-2-phenyl-benzoimidazol-1-yloxy)-acetic acid	**
AI		1-Benzyloxy-6-nitro-2-phenyl-1H-benzoimidazole	**
AJ		1-(4-Methyl-benzyloxy)-6-nitro-2-phenyl-1H-benzoimidazole	*
AK		6,8-Dimethyl-2-(4-nitro-phenyl)-5-phenyl-5H,6H-1-oxa-3,5,6,8-tetraaza-cyclopenta[a]naphthalene-4,7,9-trione	**

AL		6,8-Dimethyl-5-phenyl-2-p-tolyl-5H,6H-1-oxa-3,5,6,8-tetraaza-cyclopenta [a]naphthalene-4,7,9-trione	*
AM		2-[3-(4-Fluoro-phenyl)-1-phenyl-1H-pyrazol-4-yl methylene]-benzo [4,5] imidazo[2,1-b]thiazol-3-one	**
AN		Cobalt 5,10,15,20-Tetra-pyridin-4-yl-porphyrine	***
AO		2-[3-(4-Fluoro-phenyl)-1-phenyl-1H-pyrazol-4-ylmethylene]-5-methyl-6-vinyl-imidazo[2,1-b]thiazol-3-one	**
AP		Cobalt 5,10,15,20-Tetra-pyridin-3-yl-porphyrine	***
AQ		Zinc 5,10,15,20-Tetra-pyridin-4-yl-porphyrine	***
AR		2-(4-isopropylphenyl)-4H-chromen-4-one	***

AS		2-(3,4-Dihydroxy-phenyl)-3,5,7-trihydroxy-chromen-4-one (luteolin)	***
AT		N-isopropyl-2-[(4-methyl-5-quinolin-6-yl-4H-1,2,4-triazol-3-yl)thio]acetamide	***
AU		4-hydroxy-6-methyl-5,6-dihydro-2H-pyrano[3,2-c]quinoline-2,5-dione	***
AV		5,7-Dihydroxy-2-(4-hydroxy-phenyl)-chromen-4-one	***
AW		2-[4-(dimethylamino)phenyl]-4H-chromen-4-one	**
AX		1-(benzyloxy)-2-phenyl-1H-imidazo[4,5-b]pyridine	**
AY		2-(benzylthio)-4-phenyl-5-(1-phenyl-1H-1,2,3,4-tetraazol-5-yl)pyrimidine	**
AZ		6-fluoro-2-phenyl-4H-chromen-4-one	**
BA		7-methoxy-2-phenyl-4H-chromen-4-one	*

BB		4-(1,3-dioxo-1,3-dihydro-2H-inden-2-yliden)-2-phenyl-6-(2-pyridinyl) tetrahydro pyrrolo [3,4-c]pyrrole-1,3(2H,3aH)-dione	*
BC		2-(2-Hydroxy-3-oxo-5-p-tolyl-2,3-dihydro-furan-2-yl)-malonamic acid ethyl ester	*
BD		2-[(6-nitro-2-phenyl-1H-1,3-benzimidazol-1-yl)oxy]acetic acid	*
BE		2-(4-fluorophenyl)-4H-chromen-4-one	*
BF		1-methoxy-2-(4-methyl phenyl)-1H-imidazo [4,5-b] pyridine	*
BG		2-(4-hydroxyphenyl)-4H-chromen-4-one	*

Example 4:

In this example, the expression of a selective marker (*e.g.*, *ccdB*) is put under the direct control of a promoter repressed by MarA (*e.g.*, *fecA*, *purA*, *guaB*). Under conditions of constitutive MarA synthesis (*e.g.*, using a constitutive *mar* (*mar*^{c}) mutant), the expression of ccdB is silent. Following inactivation of MarA, the synthesis of ccdB results in cell death.

Example 5:

In this example, the expression of a selective marker (*e.g.*, *URA3*) is put under the direct control of a *promoter* repressed by MarA (*e.g.*, *fecA*, *purA*, *guaB*). Under conditions of constitutive MarA synthesis (*e.g.*, using a constitutive *mar* (*mar*^{c}) mutant), the expression of URA3 is silent. Following inactivation of MarA, and in the presence of 5-FOA the synthesis of URA3 results in cell death.

Example 6:

In this example, a purine or guanine heterotroph is constructed by inactivation of either of the chromosomal *guaB* or *purA* genes in E. coli. The *guaB* or *purA* gene is then placed into a suitable vector under the control of its natural promoter and transformed into the heterotrophic host.

Example 7: *E. coli* Biofilm Assay

The biofilm assay screens test compounds for their ability to inhibit bacteria from forming a biofilm.

Materials:

The M9 media ("M9") contained M9, casamino acids , and glucose. The test compound was dissolved in 10mg/mL DMSO stock solution.

Method:

Preparation of Inoculum

Inoculum was started the day of the experiment by adding a colony or glycerol stock stab to 2mL M9. The tube was placed in the 37 °C shaker incubator for approximately 4-6 hours. This inoculum was referred to as the "Starter inoculum." The inoculum was then removed from the shaker incubator and diluted to 1 X 10^6 cells/mL in M9.

Preparation of Controls

Typically, there were eight of each control, including a positive and negative control. For both the positive and negative controls, 2.5 µL of DMSO was added to 200 µL of M9. 25 µL of the diluted DMSO was added to 50 µL of M9 in the assay plate.

Preparation of Test Compounds

The test compounds were screened at 20 µg/mL. 2.5 µL of the test compound were taken from a plate containing 10mg/mL stock and added to 200µL of M9 and mixed. 25 µL of the diluted test compound was added to 50 µL of M9 in the assay plate. The resulting concentration of the test compound was 40 µg/mL

Preparation of Plate

75 µL of the inoculum at 1 X 10^6 cells/mL was added to each well containing compound and the positive controls. 75 µL M9 was added to the negative

controls. The final concentration of the test compound was 20 μg/mL and the final concentration of the inoculum was 2 X 10^5 cells/mL. The plates were then placed in a microplate reader (Wallac Victor^{2}V) and read OD_{535} ("Initial growth reading"). The plates were then placed in an incubator overnight at 35 °C.

In the morning, the plates were read in a microplate reader at OD_{535} ("Final growth reading.")

Addition of Crystal Violet

The inoculum was then removed from the wells and the plates were washed several times with tap water. 150 μL of Crystal Violet (0.02% Crystal Violet dissolved in water) was then added to each well.

Addition of Ethanol

The crystal violet was then removed and the plates were washed several times with tap water. 150 μL of ethanol were then added to each well, after mixing. The plates were then placed in a microplate reader and read the OD_{535}. This was referred to as the "Crystal Violet" reading.

Data Analysis

To determine whether a test compound inhibited growth, the Initial growth reading was subtracted from the Final growth reading ("Subtracted Growth").The same was done for the positive controls and averaged. The % inhibition of growth was determined by the following formula:

100-(100*Subtracted growth of sample/Average growth of Pos Controls)

To determine whether a test compound inhibited Biofilm formation, the %Inhibition of Biofilm Formation was determined using the following formula:

100-(100*Crystal Violet read of sample/Average crystal violet read of Pos Controls)

The results from the Crystal Violet assay are summarized in Table 5. In Table 5, ND indicates that a given compound did not inhibit biofilm formation in the CV assay. * indicates that the test compound inhibited some biofilm formation and ** indicates that the compound inhibited the formation of a biofilm well.

Example 8: LANCE Screening Assay for Inhibitors of MarA, SoxS, or Rob DNA-binding

This example describes a method for the identification of test compounds that inhibit the interactions of purified transcription factor such as MarA, SoxS and/or Rob with a target DNA sequence in an *in vitro* system. Such molecules will hopefully be able to inhibit this interaction *in vivo*, leading to inhibition of transcriptional regulation by these factors and ultimately in inhibition of the drug resistance and/or virulence phenotypes associated with MarA, SoxS and Rob.

Materials

The 6His-tagged MarA, SoxS and Rob purified according to respective protocol. The N-term-biotinylated double-stranded DNA has a sequence of CCG ATT TAG CAA AAC GTG GCA TCG GTC (SEQ ID NO. 5). The antibody used was the LANCE Eu-labeled anti-6xHis Antibody (Eu-αHis) (Perkin Elmer cat # AD0110) which had at least 6 Europium molecules per antibody. Streptavidin conjugated to SureLight™-Allophycocyanin (SA-APC) was obtained from Perkin Elmer (cat # CR130-100). The Assay buffer contained 20mM Hepes pH 7.6, 1mM EDTA, 10mM $(NH_4)_2SO_4$, and 30mM KCl, and 0.2% Tween-20.

Method

The plates or vials of the compounds to be tested were thawed. These stocks were at a concentration of 10mg/ml in DMSO. The solutions were allowed to thaw completely, and the plates were briefly shaken on the Titermix to redissolve any precipitated compound. Thawed aliquots of MarA, SoxS and Rob protein from the stock stored at –80°C and 1M stock of dithiothreitol stored at -20°C were then placed on ice.

Dilutions at 1:100 of the compounds were made into a fresh 96-well polystyrene plate. The dilutions were prepared with 100% DMSO to give a final concentration of 100 µg/ml solutions. The dilutions were vortexed on a Titermix.

Fresh DTT was added to 25-50 mL of assay buffer to produce a 1mM final concentration. Next, 90 µl of assay buffer was added to each of the 10 µl protein aliquots, and the solution was mixed by pipetting. These proteins were diluted to give the required amount of each of the diluted proteins, resulting in 20 µl of diluted protein per well. In preparing the solutions, 20% excess was made to allow enough for control wells. Typically, depending on the protein preps and the initial binding curves that were performed, 1000-2000 fmoles of each protein was required per well. The diluted protein solutions were the placed on ice.

Three tests plates per plate of compound (for MarA, SoxS and Rob) were prepared. Using a multichannel pipet, 5µl of the compound was added to each well. 5µl of DMSO was added to the blank and control wells, and 5µl of the control inhibitor was added to the respective wells.

Using the multichannel pipet, 20µl of protein was added to all wells except those designated "blank". To these blank wells, 20µl of assay buffer was added. The plates were covered with a plate sealer and incubated at room temperature, shaking on the Titermix, for 30minutes.

Next, the DNA solution was prepared, with enough for at least 20% more wells than were tested. 15µl (0.4 fmoles) was added per well. Then the DNA was diluted in assay buffer, and vortexed briefly to mix. The plate sealer was removed, and 15 µl of DNA solution was added to all of the wells. the plates were then resealed, and returned to the Titermix for a further 30 minutes.

After 25 minutes, the antibody solution was prepared. 0.4 fmoles of SA-APC and 0.125 fmoles of Eu-αHis were added per well in a total volume of 10µl. Amounts were prepared sufficient for at least 20% excess. The plate sealer was the removed and 10 µl of antibody solution was added to every well. The plates were subsequently resealed, placed on the Titermix, and covered with aluminum foil. The plates were mixed for 1 hour. The plates were then read on the Wallac Victor V, using the LANCE 615/665 protocol.

Data processing

For each plate, the mean control (i.e. signal from protein and DNA without inhibition), mean blank (background signal without protein) and mean inhibitor (P001407) $LANCE_{665}$ counts were determined. The percentage inhibition by each molecule (each test well) was then determined according to the following equation:

% Inhibition = 100-(((test–mean blank)/(mean control-mean blank)*100)

Compounds that gave 40% or greater inhibition were identified as hits and screened again for IC50.

IC_{50} screening

The protocol used was identical to that outlined above, except that only 10 compounds were assayed per plate. The testing concentrations started at 10 µg/ml and were diluted two-fold from 10 to 0.078µg/ml.

IC_{50} Data processing

Percent inhibition was calculated as shown above. Percent inhibition was then plotted vs. log (conc. Inhibitor) using Graph pad Prism software. The IC_{50} concentration was determined as the concentration that gives 50% inhibition.

The data from this example is also summarized in Table 5. *** indicates that a particular test compound inhibited the particular bacteria very well, ** indicates that the particular test compound inhibited the particular bacteria well, and * indicates that the particular test compound inhibited the particular bacteria to some extent.

Example 9: Luciferase Assay

The luciferase assay is used to determine if any of the compounds tested reduce the luminescent signal. This indicates that the test compounds affect regulation of micF, which in turn is regulated by Mar.

Materials

The bacteria used were *E.coli* AG112KmicF-Luc. The negative control Bacteria were *E.coli* AG112. The test compounds were prepared in a 10mg/mL DMSO stock solution.

Methods

Preparation of Inoculum

Inoculum (or "Starter Inoculum") was started the night before the day of the experiment by adding either a colony or stab of a glycerol stock to 2mL of LB Broth. The Starter Inoculum was then placed in a 37 °C shaker incubator and allowed to grow overnight.

The following day, the Starter Inoculum was removed from the shaker and added to fresh LB Broth. For each plate to be assayed, 6mL of LB broth was prepared, with 5-10μL of Starter Inoculum being added per mL of added LB to form the "Test Inoculum". Typically, four plates of test compounds were assayed. In this typical example, 6mL of LB Broth was used for each plate, or 24 mL of LB, followed by the addition of 5μL/mL of Starter Inoculum, or 120μL of Starter Inoculum to form the Test Inoculum.

Following preparation of the Test Inoculum, the Test Inoculum was placed in a 37 degree Celsius shaker and incubated for about 4 hours. The Test Inoculum was monitored for bacterial growth by taking OD readings at 535 nm on a spectrophotometer. The Test inoculum should be removed when the OD reaches between 0.6 and 1.5.

Preparation of Controls

Positive and negative controls were created by adding 2uL DMSO to 198uL LB Broth. At least 4 of each control were generated. Typically, there were 8 of each. 50uL of diluted DMSO was added to 50uL LB Broth in the assay plates.

Preparation of Compounds

The compounds were screened at 25ug/mL. Two identical plates of each compound were set up: 1 clear plate for growth (or "Clear Plate"), 1 white plate for luminescence (or "White Plate"). Next, 2 μL of each compound was taken from the daughter plate (containing 10mg/mL stock) and added to 198 μL of LB Broth. The sample was then mixed. Next, 25 μL of the diluted test compound was added to 25 μL of LB Broth in all of the assay plates. The concentration of the compound at this stage was 50 μg/mL.

Preparation of Plate

50 μL of the Test Inoculum was added to each well of the plates, except for the negative controls. Half of the negative controls received 50 μL of AG112, while the other half of the negative controls received 50 μL LB Broth. The final concentration of the test compound was 25 μg/mL.

The Clear Plates were placed in the plate reader and read at OD_{535}. This was the "Initial" growth read. The plates were then incubated plates for 5 hours at 37 degrees Celsius. After 5 hours, the plates were removed from the incubator. The Clear Plates were placed in the plate reader and read at OD_{535}. This was the "Final" growth read.

100 μL of Promega Steady-Glo reagent was added to each well (including all controls) in the White Plates. The plates, covered with aluminum foil, were then shaken on a plate shaker set at 10000 rpm for 10 min. The plates were then placed in plate reader and read on luminescence for 1sec per well. This was the LUMINESCENT read.

Data Analysis

To determine whether the test compound inhibited growth, the Initial growth read was subtracted from the Final growth read. This was the Subtracted Growth. The same calculation was performed for the positive controls. The results for the positive controls were averaged. The %Inhibition of Growth was determined using the following formula:

100-(100*Subtracted growth of sample/Average growth of Pos Controls)

To determine whether compound inhibits Luciferase, use the following equation:

100-(100*Luminescence for Compound/Average Luminescence of Pos Controls)

ND indicates that a particular test compound did not appear to decrease the lumninesce of in this particular assay. * indicates that the luminescence was decreased somewhat and ** indicates that the luminescence was decreased a substantial amount. The results from this assay are also shown in Table 5.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, numerous equivalents to the specific polypeptides, nucleic acids, methods, assays and reagents described herein. Such equivalents are considered to be within the scope of this invention and are covered by the following claims.

TABLE 5

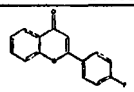
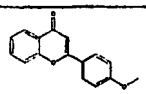
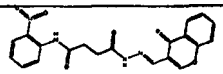
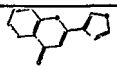
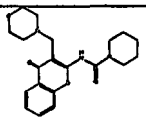
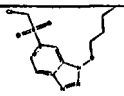
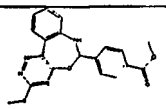
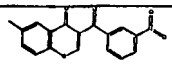
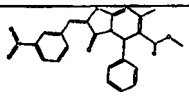
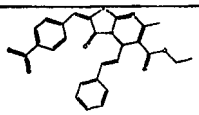
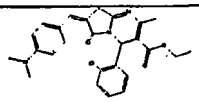
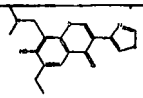
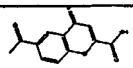
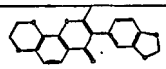
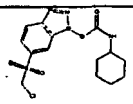
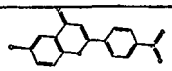
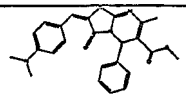
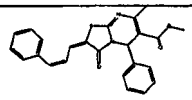
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
BH		**	*	**	**	**	**
BI		**	NT	NT	**	**	ND
BJ		***	***	*	**	ND	**
BK		*	*	**	**	ND	ND
BL		*	**	*	**	**	ND
BN		*	*	*	**	**	ND
BO		*	**	*	**	**	ND
BP		*	NT	NT	**	*	**
BQ		*	*	***	**	ND	**
BR		***	***	***	**	ND	ND
BS		*	*	*	**	**	**
BT		*	**	*	**	*	*
BU		*	NT	NT	**	ND	ND
BV		*	NT	NT	**	**	**
BW		NT	***	**	**	ND	ND
BX		*	NT	NT	**	ND	ND
BY		*	*	***	**	ND	**
BZ		***	*	**	**	ND	ND

TABLE 5

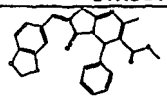
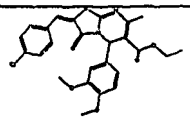
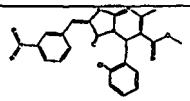
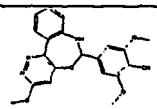
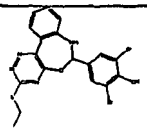
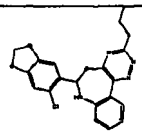
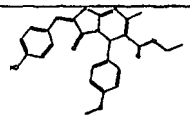
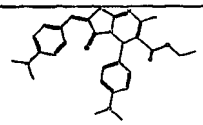
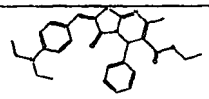
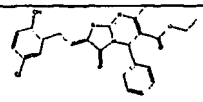
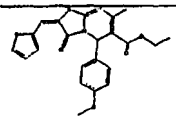
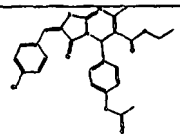
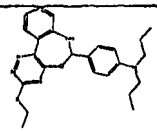
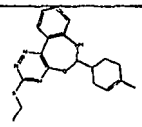
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
CA		***	**	NT	**	ND	**
CB		*	*	*	**	ND	**
CC		*	*	**	**	*	ND
CD		NT	***	***	**	ND	**
CE		**	**	*	**	*	ND
CF		*	NT	NT	**	ND	**
CG		*	NT	NT	**	ND	**
CH		*	NT	NT	**	ND	ND
CI		*	NT	NT	**	ND	**
CJ		**	*	*	**	ND	ND
CK		**	***	**	**	ND	ND
CL		**	*	*	**	ND	ND
CM		*	*	*	**	**	**
CN		***	***	*	**	**	ND

TABLE 5

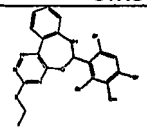
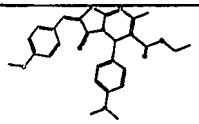
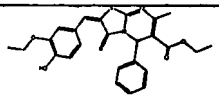
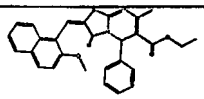
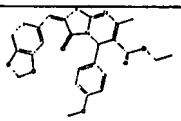
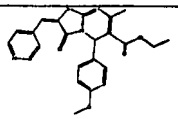
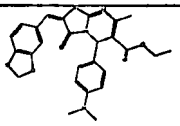
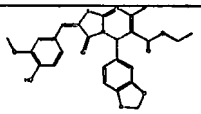
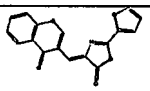
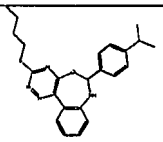
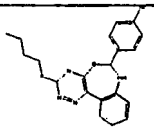
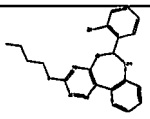
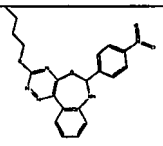
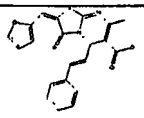
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lac
CO		*	*	**	**	**	*
CP		*	*	*	**	**	*
CQ		*	*	*	**	**	*
CR		*	**	***	**	**	ND
CS		***	***	*	**	**	ND
CT		NT	***	***	**	ND	*
CU		NT	**	*	**	ND	ND
CV		*	NT	NT	**	ND	ND
CW		*	*	***	**	ND	**
CX		*	***	***	**	ND	**
CY		*	*	*	**	ND	ND
CZ		*	*	**	**	ND	ND
DA		*	*	*	**	ND	ND
DB		**	*	*	**	ND	ND

TABLE 5

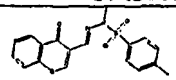
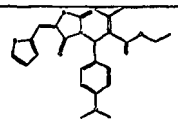
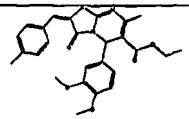
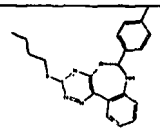
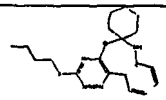
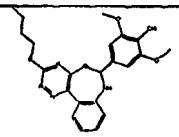
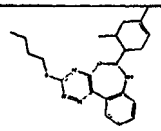
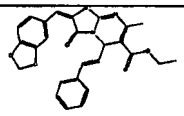
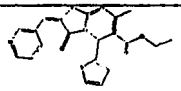
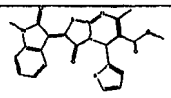
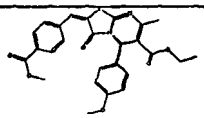
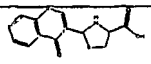
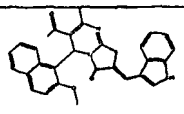
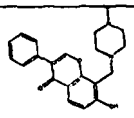
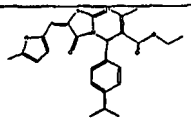
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
DC		*	*	*	**	ND	*
DD		*	*	*	**	ND	*
DE		*	*	**	**	**	ND
DF		*	*	*	**	ND	**
DG		*	*	*	**	**	**
DH		*	*	**	**	**	ND
DI		*	*	*	**	**	**
DJ		**	NT	NT	**	ND	ND
DK		*	NT	NT	**	ND	ND
DL		***	*	***	**	ND	**
DM		***	*	*	**	ND	ND
DN		*	*	*	**	ND	ND
DO		*	*	*	**	ND	*
DP		***	*	*	**	ND	ND
DQ		*	*	*	**	ND	ND

TABLE 5

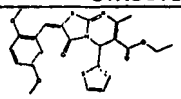
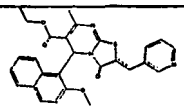
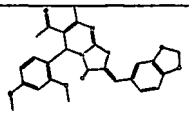
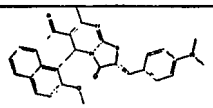
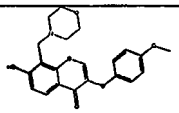
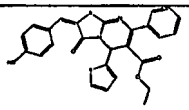
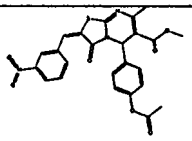
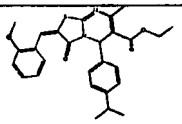
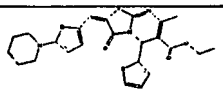
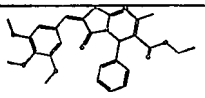
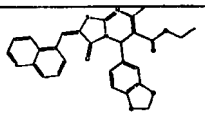
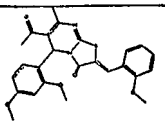
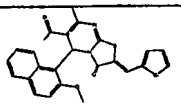
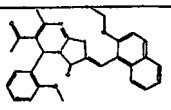
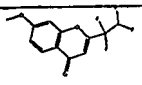
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	DD	CV	Luc
DR		***	*	*	**	ND	ND
DS		*	*	*	**	ND	*
DT		*	*	*	**	ND	*
DU		***	***	*	**	ND	**
DV		***	***	**	**	ND	**
DW		**	*	**	**	ND	**
DX		***	***	**	**	ND	**
DY		*	*	**	**	ND	**
DZ		*	*	**	**	**	**
EA		*	NT	NT	**	*	*
EB		*	NT	NT	**	*	**
EC		*	NT	NT	**	**	ND
ED		***	***	*	**	**	**
EF		**	***	***	**	*	**
EG		*	**	*	**	**	**

TABLE 5

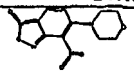
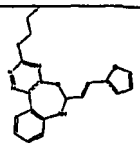
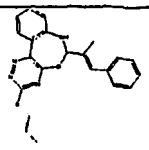
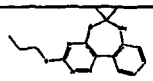
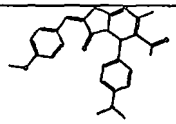
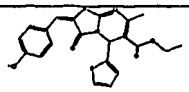
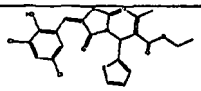
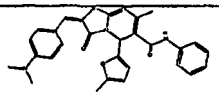
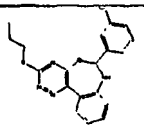
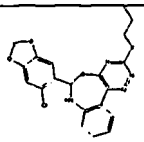
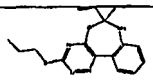
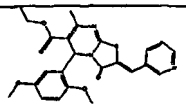
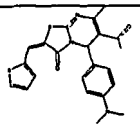
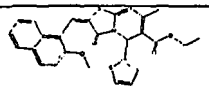
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	GD	CV	Luc
EH		***	***	**	**	*	**
EI		*	NT	NT	**	ND	ND
EJ		*	NT	NT	**	ND	ND
EK		***	NT	NT	**	ND	*
EL		*	NT	NT	**	ND	ND
EM		**	NT	NT	**	ND	ND
EN		*	NT	NT	**	ND	*
EO		*	*	*	**	ND	**
EQ		*	*	*	**	ND	ND
ER		*	*	*	**	ND	**
ES		*	*	***	**	ND	ND
ET		**	*	*	**	ND	**
EU		*	*	*	**	ND	ND
EV		***	*	*	**	ND	ND
EW		*	*	**	**	ND	**

TABLE 5

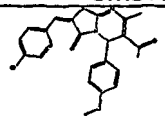
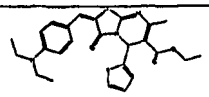
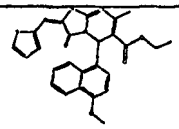
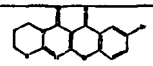
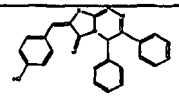
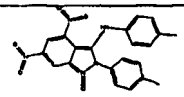
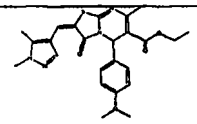
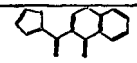
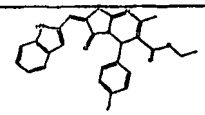
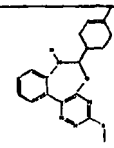
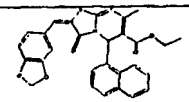
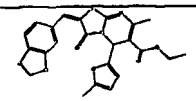
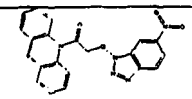
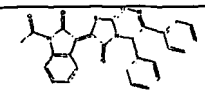
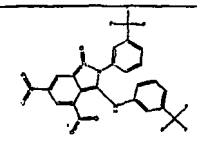
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
EX		*	*	**	**	ND	ND
EY		*	*	*	**	ND	*
EZ		*	*	*	**	ND	*
FA		*	*	*	**	ND	*
FB		*	*	*	**	ND	*
FC		*	*	*	**	*	*
FD		*	*	*	**	*	**
FE		*	**	*	**	**	ND
FG		**	***	**	**	*	ND
FH		*	*	*	**	ND	**
FI		*	*	*	**	ND	ND
FJ		**	**	*	**	ND	**
FK		***	*	***	**	ND	**
FL		*	*	**	**	ND	**
FM		*	*	*	**	ND	**

TABLE 5

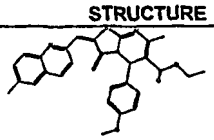
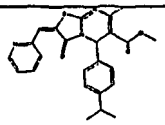
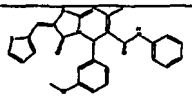
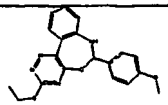
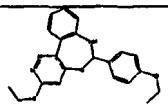
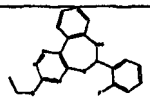
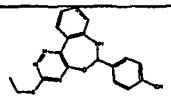
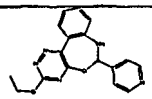
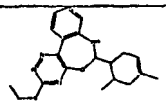
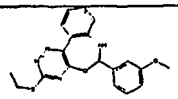
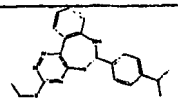
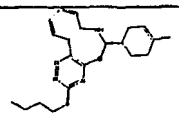
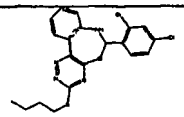
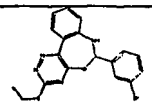
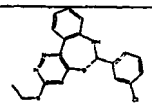
ID	STRUCTURE	MarA Lance	BoxB Lance	Rob Lance	OD	CV	Luc
FN		***	**	NT	**	ND	*
FO		***	**	*	**	*	*
FP		***	***	*	**	**	**
FQ		*	**	**	**	**	*
FR		*	NT	NT	**	*	**
FS		*	**	*	**	**	ND
FT		NT	*	*	**	ND	**
FU		**	*	***	**	ND	**
FV		*	*	*	**	ND	**
FW		*	*	*	**	ND	ND
FX		*	*	*	**	ND	**
FY		*	*	*	**	ND	**
FZ		*	*	*	**	ND	**
GA		*	*	**	**	ND	ND
GB		*	*	*	**	ND	*

TABLE 5

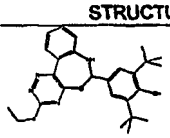
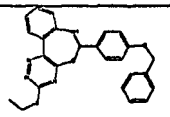
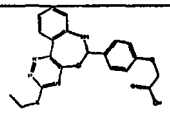
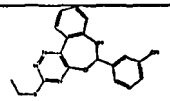
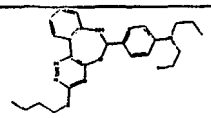
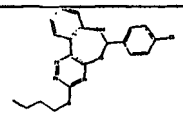
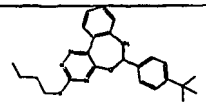
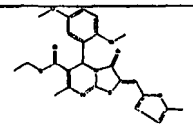
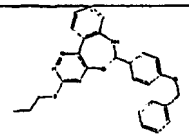
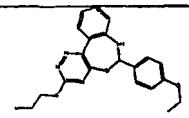
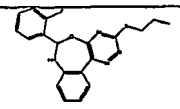
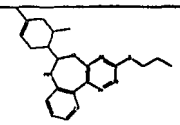
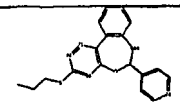
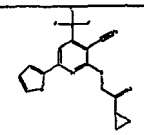
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
GC		*	*	*	**	ND	**
GD		*	*	*	**	ND	ND
GE		*	*	*	**	ND	ND
GF		**	*	*	**	ND	ND
GG		*	*	*	**	ND	*
GH		*	*	*	**	ND	ND
GI		**	**	***	**	ND	**
GJ		*	*	**	**	ND	ND
GK		*	*	*	**	ND	*
GL		*	*	*	**	ND	ND
GM		*	*	***	**	ND	ND
GN		*	*	*	**	ND	ND
GO		*	*	*	**	ND	*
GP		*	**	**	**	**	**

TABLE 5

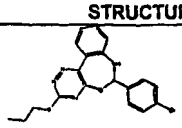
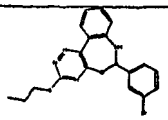
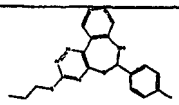
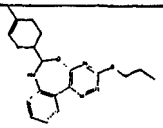
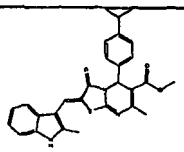
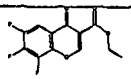
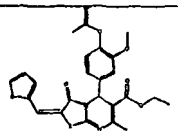
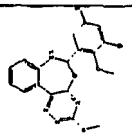
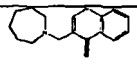
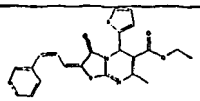
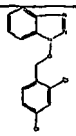
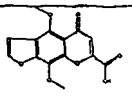
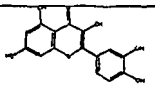
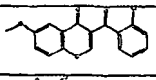
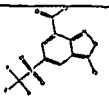
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	CD	CV	Luc
GQ		**	***	**	**	*	**
GR		*	***	***	**	*	ND
GS		***	***	***	**	**	ND
GT		*	**	**	**	*	ND
GU		**	***	*	**	**	ND
GV		*	*	*	**	ND	ND
GW		*	*	**	**	ND	**
GX		*	*	***	**	ND	**
GY		*	*	*	**	ND	**
GZ		**	*	**	**	ND	**
HA		*	*	*	**	ND	**
HB		*	*	*	**	ND	*
HC		***	*	NT	**	ND	**
HD		**	*	***	**	ND	*
HE		*	*	*	**	**	**

TABLE 5

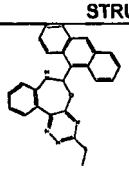
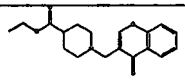
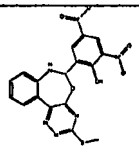
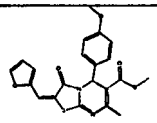
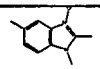
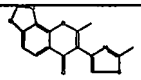
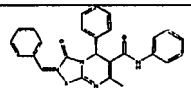
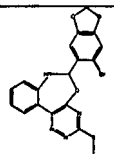
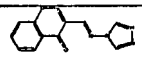
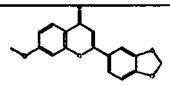
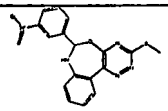
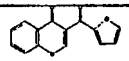
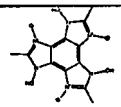
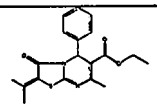
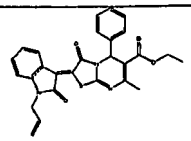
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
HF		**	*	*	**	**	ND
HG		*	*	*	**	**	**
HH		*	*	*	**	**	**
HI		*	*	*	**	**	**
HJ		*	**	*	**	**	**
HK		*	**	**	**	**	*
HL		***	***	*	**	**	**
HM		NT	**	*	**	ND	**
HN		NT	**	**	**	ND	ND
HO		NT	*	**	**	ND	ND
HP		NT	*	*	**	ND	*
HQ		*	**	*	**	**	**
HR		*	*	*	**	**	**
HS		**	NT	NT	**	ND	*
HT		*	NT	NT	**	ND	**
HU		***	***	*	**	ND	**

TABLE 5

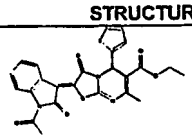
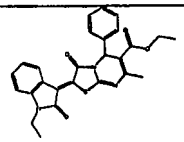
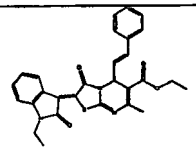
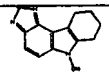
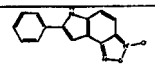
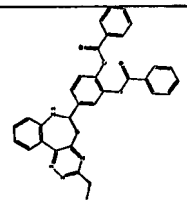
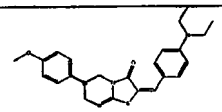
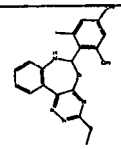
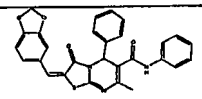
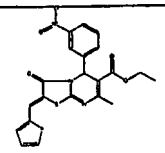
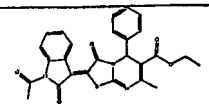
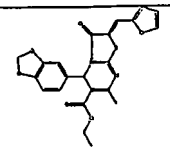
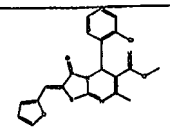
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lac
HV		*	*	*	**	ND	**
HW		*	***	*	**	ND	**
HX		*	*	*	**	ND	**
HY		*	*	*	**	ND	*
HZ		*	*	*	**	ND	**
IA		*	*	*	**	ND	*
IB		*	*	*	**	ND	ND
IC		*	*	*	**	ND	ND
ID		*	*	*	**	ND	*
IE		***	***	*	**	ND	**
IF		*	*	*	**	ND	ND
IG		*	*	*	**	ND	ND
IH		*	*	*	**	ND	*

TABLE 5

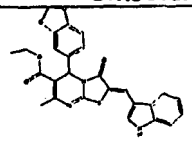
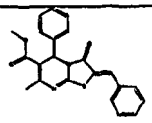
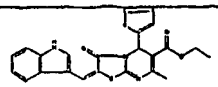
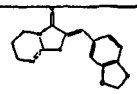
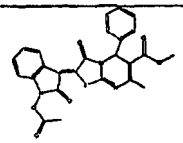
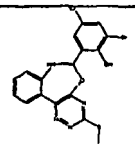
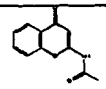
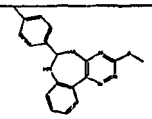
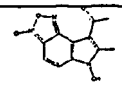
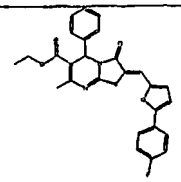
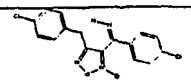
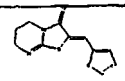
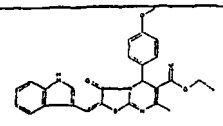
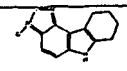
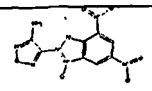
ID	STRUCTURE	MarA Lance	BoxS Lance	Rob Lance	OD	CV	Luc
II		**	*	*	**	ND	**
IJ		**	*	*	**	ND	*
IK		*	*	*	**	ND	ND
IL		***	*	*	**	ND	**
IM		*	*	*	**	ND	**
IN		*	*	*	**	ND	ND
IO		*	*	*	**	ND	ND
IP		*	*	*	**	ND	ND
IQ		*	*	*	**	ND	ND
IR		*	*	**	**	ND	**
IS		*	***	**	**	ND	*
IT		*	*	*	**	ND	*
IU		***	*	**	**	ND	ND
IV		*	*	*	**	ND	ND
IW		*	*	*	**	ND	ND

TABLE 5

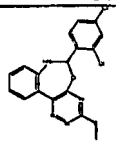
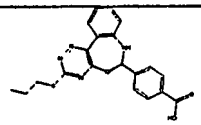
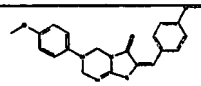
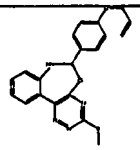
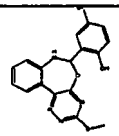
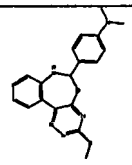
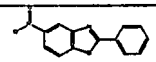
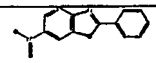
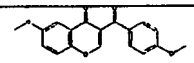
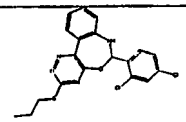
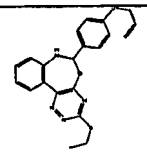
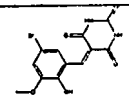
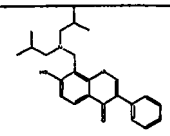
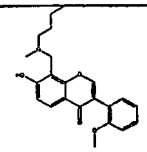
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
IX		**	*	*	**	ND	*
IY		*	*	*	**	ND	**
IZ		*	*	*	**	ND	*
JA		*	*	*	**	ND	*
JB		**	*	*	**	ND	ND
JD		*	*	*	**	ND	*
JE		*	*	*	**	**	*
JF		**	**	***	**	*	*
JG		*	*	**	**	**	ND
JH		*	*	*	**	**	**
JI		*	*	**	**	**	**
JJ		**	*	**	**	ND	**
JK		**	*	*	**	ND	**
JL		**	*	*	**	ND	ND

TABLE 5

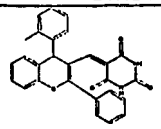
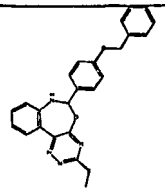
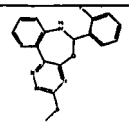
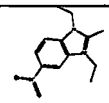
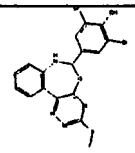
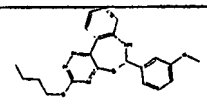
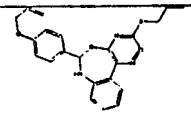
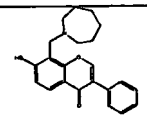
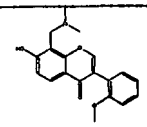
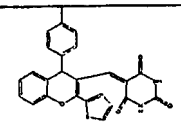
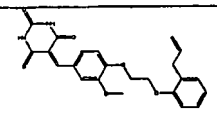
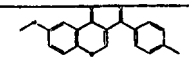
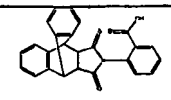
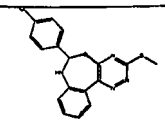
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
JM		*	*	*	**	ND	**
JN		***	*	NT	**	ND	**
JO		***	*	NT	**	ND	**
JP		*	*	*	**	*	*
JQ		*	*	*	**	**	**
JR		***	***	***	**	**	ND
JZ		**	*	*	**	**	**
KA		*	*	*	**	**	**
KB		*	*	*	**	**	**
KC		*	*	*	**	**	**
KD		*	NT	NT	**	*	**
KE		*	NT	NT	**	**	*
KF		***	***	**	**	**	**
KG		NT	*	*	**	ND	*

TABLE 5

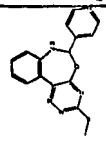
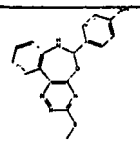
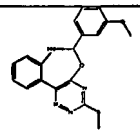
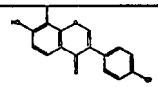
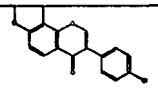
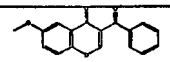
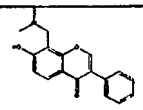
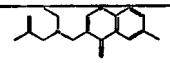
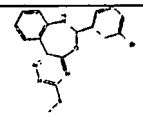
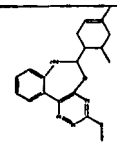
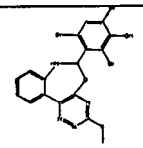
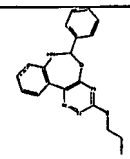
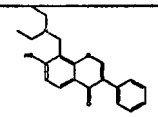
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
KH		NT	*	***	**	ND	**
KI		NT	*	*	**	ND	ND
KJ		NT	**	***	**	ND	ND
KK		*	NT	NT	**	ND	**
KL		*	*	*	**	ND	ND
KM		*	*	**	**	ND	ND
KN		***	*	*	**	ND	**
KO		*	***	***	**	ND	**
KP		*	*	**	**	ND	**
KQ		*	*	*	**	ND	*
KR		***	**	*	**	ND	ND
KS		*	*	*	**	ND	*
KT		*	*	*	**	ND	*
KU		*	*	**	**	ND	**

TABLE 5

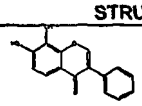
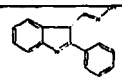
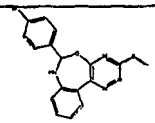
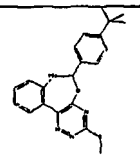
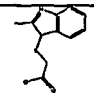
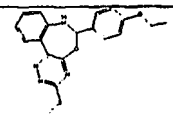
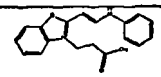
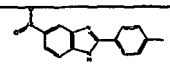
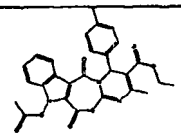
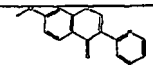
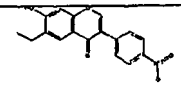
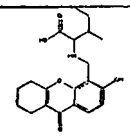
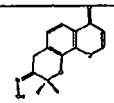
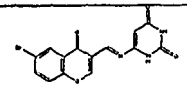
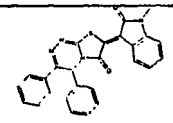
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
KV		*	*	*	**	ND	*
KW		*	*	*	**	ND	ND
KX		*	**	*	**	ND	ND
KY		*	**	*	**	ND	**
KZ		**	*	**	**	ND	ND
LA		*	*	*	**	ND	ND
LB		*	*	*	**	ND	ND
LC		*	*	*	**	ND	ND
LD		*	*	*	**	ND	*
LE		*	*	*	**	ND	ND
LF		*	*	**	**	ND	ND
LG		*	*	*	**	ND	*
LH		**	*	*	**	ND	*
LI		*	*	*	**	ND	*
LJ		*	**	*	**	*	**
LK		*	*	*	**	**	*

TABLE 5

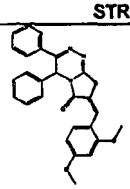
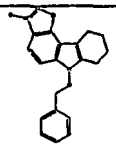
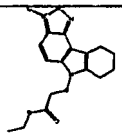
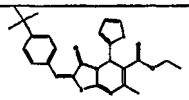
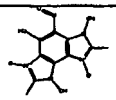
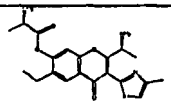
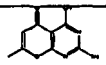
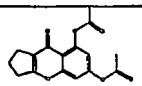
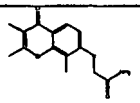
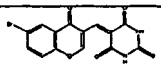
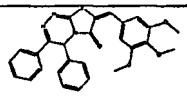
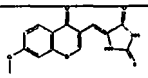
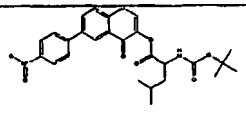
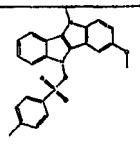
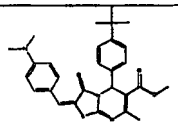
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
LL		•	•	•	••	••	••
LM		••	•••	•	••	••	ND
LN		••	•	•	••	ND	••
LO		•••	•••	••	••	ND	••
LP		•••	•	•	••	ND	•
LQ		••	•	•••	••	ND	•
LR		•	•	•	••	ND	••
LS		••	••	•	••	ND	•
LT		••	••	••	••	ND	ND
LU		•••	•••	NT	••	ND	ND
LV		•	•	••	••	••	•
LW		•	•	•	••	••	••
LX		•••	•	•	••	••	••
LY		•	•	•	••	••	••
LZ		•	•	•	••	••	••

TABLE 5

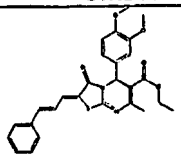
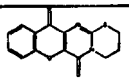
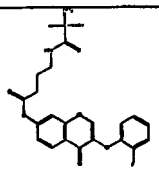
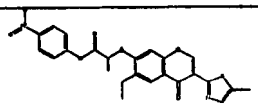
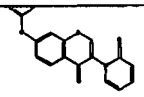
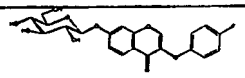
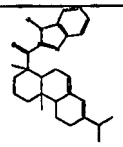
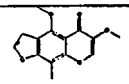
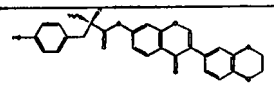
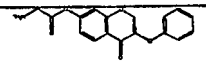
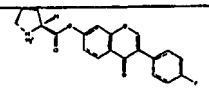
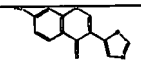
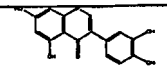
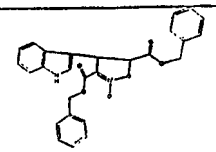
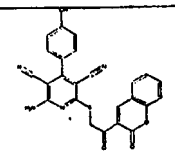
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lux
MA		*	NT	NT	**	*	*
MB		*	NT	NT	**	**	ND
MC		*	**	*	**	*	*
MD		***	***	**	**	**	*
ME		***	***	**	**	**	*
MF		***	***	***	**	**	ND
MG		*	**	**	**	**	*
MH		***	***	***	**	**	*
MI		NT	*	*	**	ND	**
MJ		NT	*	*	**	ND	*
MK		NT	**	**	**	ND	ND
ML		NT	**	***	**	ND	**
MM		NT	***	***	**	ND	ND
MN		*	**	*	**	*	ND
MO		*	*	*	**	**	**

TABLE 5

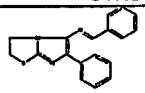
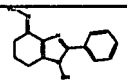
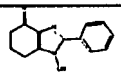
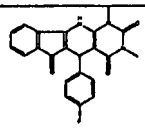
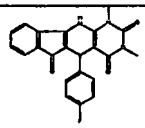
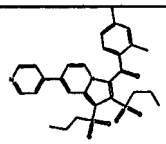
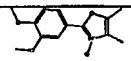
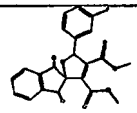
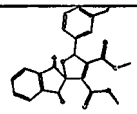
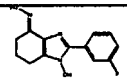
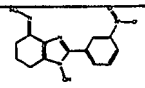
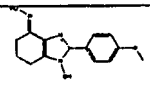
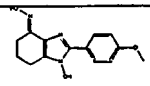
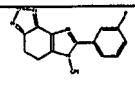
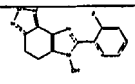
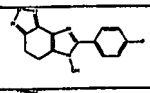
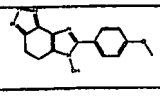
ID	STRUCTURE	MarA Lance	Sox5 Lance	Rob Lance	OD	CV	Luc
MP		*	**	*	**	**	**
MQ		*	***	*	**	**	ND
MR		**	NT	NT	**	ND	**
MS		*	NT	NT	**	ND	ND
MT		*	NT	NT	**	ND	ND
MU		**	*	*	**	ND	*
MV		*	*	*	**	ND	**
MX		***	***	**	**	ND	ND
MY		***	*	*	**	ND	**
MZ		*	*	*	**	ND	*
NA		***	*	*	**	ND	*
NB		**	*	**	**	ND	*
NC		*	*	*	**	ND	*
ND		***	*	*	**	ND	ND
NE		*	*	*	**	ND	**
NF		***	**	*	**	ND	**
NG		*	*	*	**	ND	**

TABLE 5

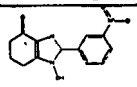
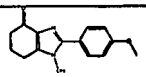
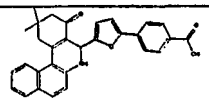
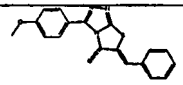
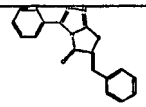
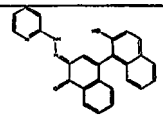
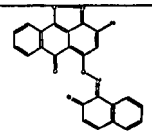
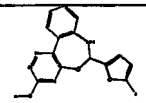
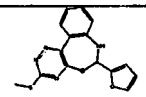
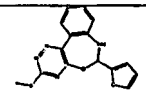
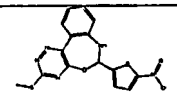
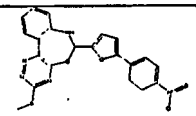
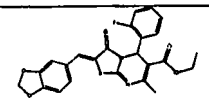
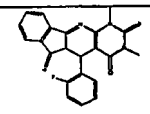
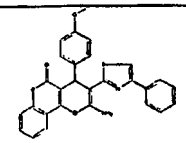
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	CD	CV	Luc
NH		*	*	*	**	ND	ND
NI		*	*	*	**	ND	ND
NJ		*	*	*	**	ND	ND
NK		*	*	*	**	ND	ND
NL		*	*	***	**	ND	*
NM		*	**	*	**	ND	**
NN		*	*	**	**	ND	*
NO		*	**	**	**	ND	*
NP		*	**	*	**	ND	ND
NQ		*	*	*	**	ND	ND
NR		*	*	*	**	ND	ND
NS		*	*	*	**	ND	ND
NT		*	*	*	**	ND	ND
NU		*	*	*	**	ND	ND
NV		*	*	*	**	ND	ND

TABLE 5

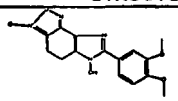
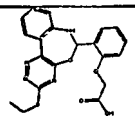
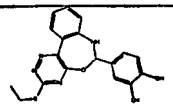
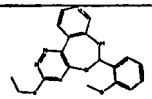
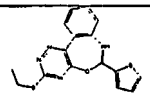
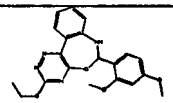
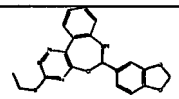
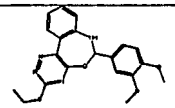
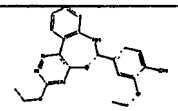
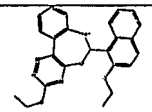
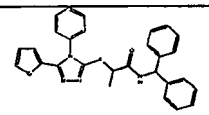
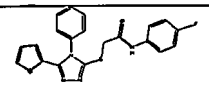
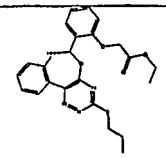
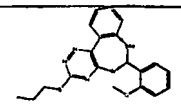
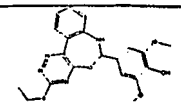
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
NW		*	*	*	**	ND	*
NX		**	*	*	**	ND	**
NY		*	*	*	**	ND	*
NZ		*	*	*	**	ND	ND
OA		***	***	*	**	ND	ND
OB		*	*	*	**	ND	**
OC		**	*	*	**	ND	ND
OD		***	*	**	**	ND	*
OE		*	*	*	**	ND	ND
OF		*	*	***	**	*	ND
OG		*	*	*	**	*	**
OH		*	*	*	**	*	**
OI		*	**	**	**	**	ND
OJ		*	**	**	**	*	ND
OK		***	***	**	**	**	ND

TABLE 5

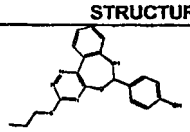
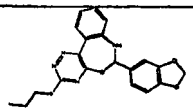
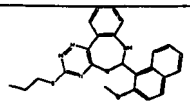
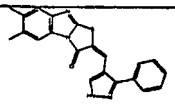
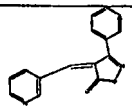
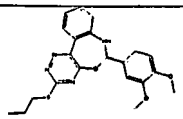
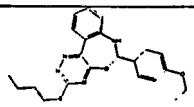
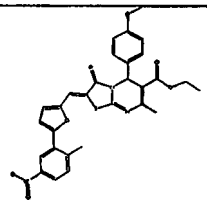
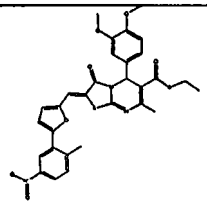
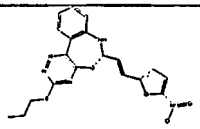
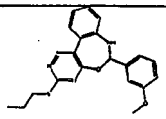
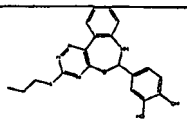
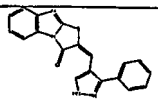
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
OL		*	**	***	**	**	*
OM		*	**	***	**	*	ND
ON		**	*	***	**	ND	**
OO		*	*	*	**	ND	*
OP		*	*	*	**	ND	**
OQ		**	*	*	**	ND	*
OR		**	*	*	**	ND	**
OS		**	*	*	**	ND	*
OT		**	**	**	**	ND	**
OU		*	**	**	**	ND	ND
OV		***	**	NT	**	ND	**
OW		*	*	**	**	**	ND
OX		*	*	*	**	**	*

TABLE 5

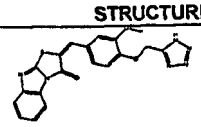
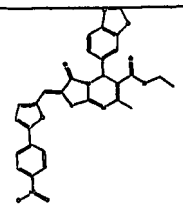
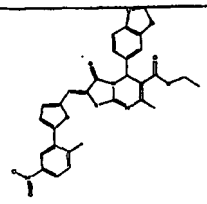
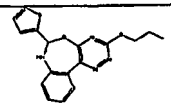
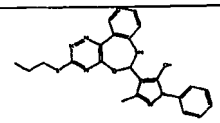
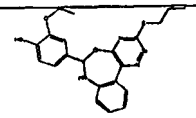
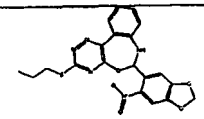
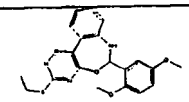
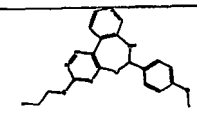
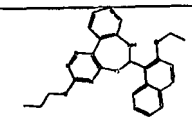
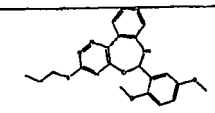
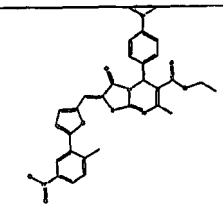
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lux
OY		*	*	*	**	**	ND
OZ		*	*	*	**	*	**
PA		***	NT	NT	**	**	*
PB		***	***	**	**	**	ND
PC		**	***	*	**	*	**
PD		NT	***	*	**	ND	ND
PD		NT	***	*	**	ND	**
PE		NT	**	*	**	ND	ND
PF		NT	***	**	**	ND	*
PG		NT	*	*	**	ND	ND
PH		NT	*	*	**	ND	ND
PI		NT	**	*	**	ND	ND

TABLE 5

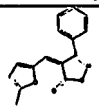
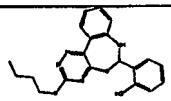
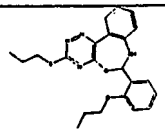
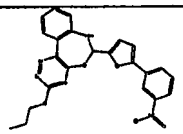
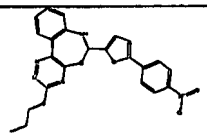
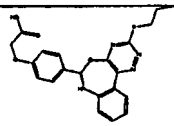
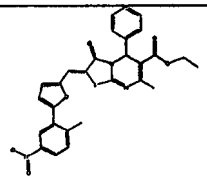
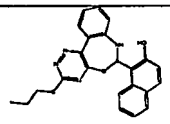
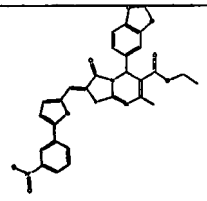
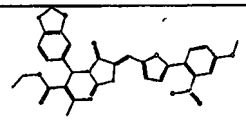
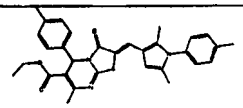
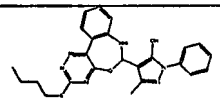
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lac
PJ		NT	*	**	**	ND	**
PK		*	**	*	**	**	**
PL		*	**	*	**	**	**
PM		*	**	*	**	*	**
PN		*	**	**	**	**	*
PO		*	*	*	**	*	ND
PP		*	NT	NT	**	ND	ND
PQ		**	NT	NT	**	ND	ND
PR		***	NT	NT	**	ND	ND
PS		*	NT	NT	**	ND	**
PT		*	*	*	**	ND	ND
PU		*	*	*	**	ND	**

TABLE 5

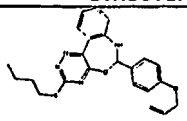
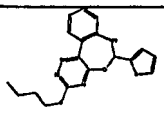
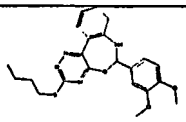
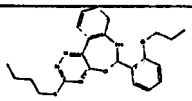
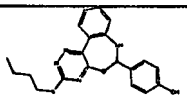
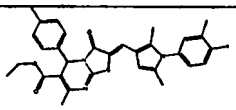
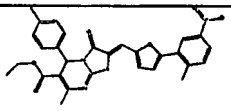
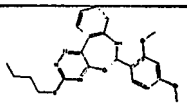
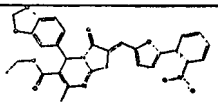
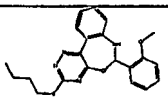
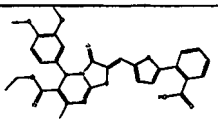
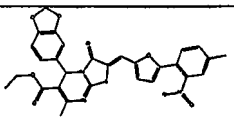
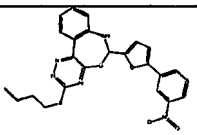
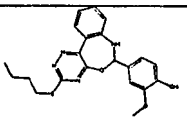
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
PV		•	•	•	••	ND	••
PW		•	•	•	••	ND	•
PX		•	•	•	••	ND	••
PY		•	•	•	••	ND	••
PZ		•	•	•	••	ND	••
QA		•	•	••	••	ND	••
QB		•••	•	••	••	ND	•
QC		•••	•	•	••	ND	ND
QD		•	•	•	••	ND	ND
QE		•	•	•	••	ND	•
QF		•••	•	•	••	ND	ND
QG		•	•	•	••	ND	••
QH		•	•	•	••	ND	•
QI		•••	•••	•••	••	ND	•

TABLE 5

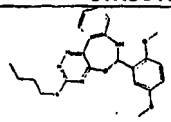
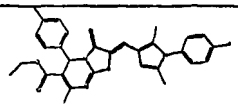
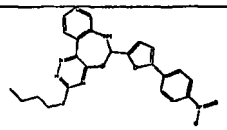
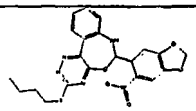
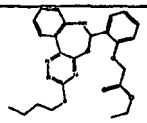
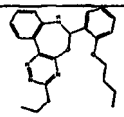
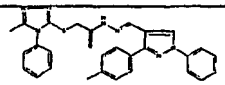
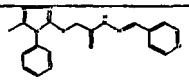
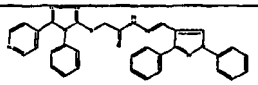
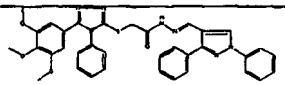
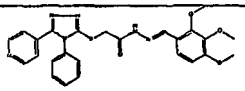
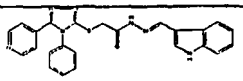
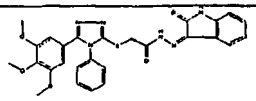
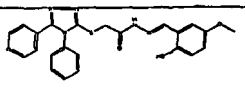
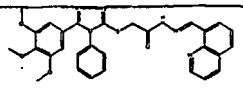
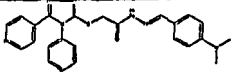
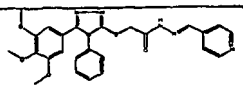
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	cp	cv	Lac
QJ		+	+	++	++	ND	+
QK		+	+	+	++	ND	++
QL		++	+	+	++	ND	ND
QM		+	+	+	++	ND	ND
QN		+	+	+	++	ND	ND
QO		+	+	+	++	ND	+
QP		+	+	+	++	ND	+
QQ		+	+	++	++	ND	ND
QR		++	+	+++	++	ND	ND
QS		+	++	+++	++	ND	ND
QT		+	+	+++	++	ND	ND
QU		++	++	++	++	ND	+
QV		+	++	++	++	ND	+
QW		+	+	++	++	ND	ND
QX		++	+	++	++	ND	ND
QY		+	+	++	++	ND	ND
QZ		++	+	+	++	ND	ND

TABLE 5

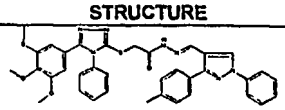
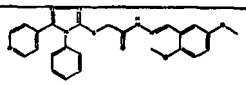
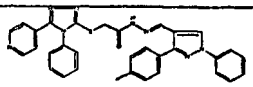
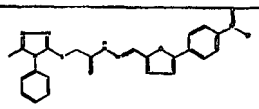
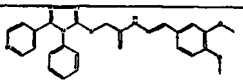
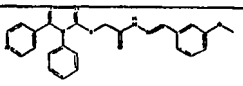
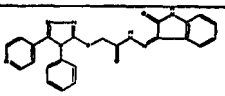
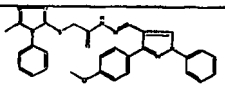
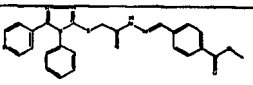
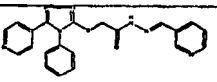
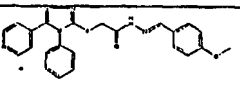
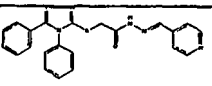
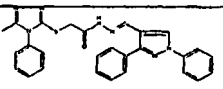
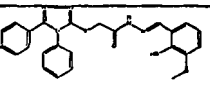
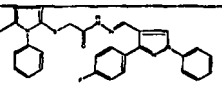
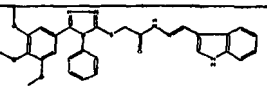
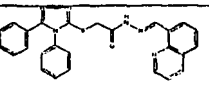
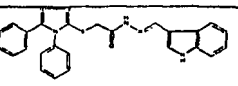
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
RA		**	*	**	**	ND	**
RB		*	*	*	**	ND	*
RC		*	**	*	**	ND	ND
RD		**	***	***	**	*	ND
RE		*	**	***	**	*	ND
RF		**	***	**	**	*	**
RG		*	*	**	**	*	**
RH		*	*	**	**	*	**
RI		*	*	*	**	*	ND
RJ		**	**	*	**	*	**
RK		*	***	**	**	*	**
RL		**	***	**	**	**	ND
RM		*	*	*	**	*	ND
RN		*	*	**	**	*	*
RO		*	*	**	**	*	*
RP		**	**	***	**	*	**
RQ		***	***	*	**	*	**
RS		***	*	***	**	ND	**

TABLE 5

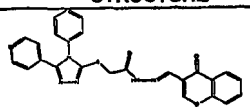
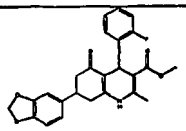
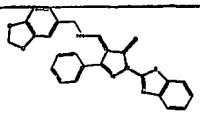
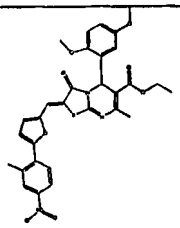
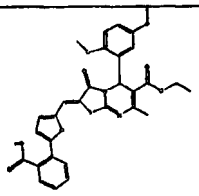
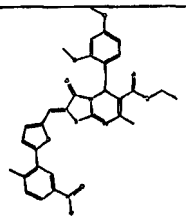
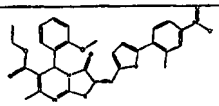
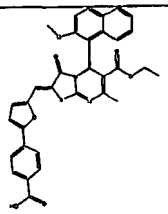
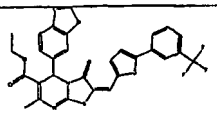
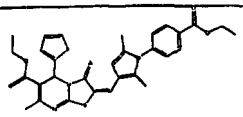
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
RT		***	***	***	**	ND	*
RU		***	**	**	**	ND	*
RV		**	*	*	**	ND	**
RW		**	**	*	**	ND	**
RX		**	***	*	**	ND	ND
RY		**	*	*	**	ND	**
RZ		*	*	*	**	ND	**
SA		*	*	*	**	ND	*
SB		**	*	*	**	ND	**
SC		***	*	*	**	ND	**

TABLE 5

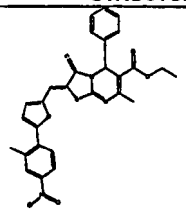
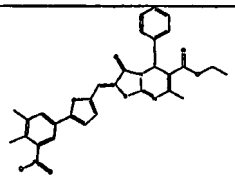
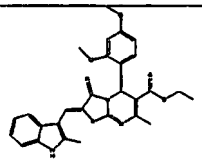
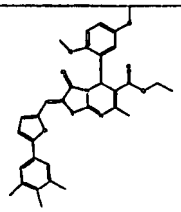
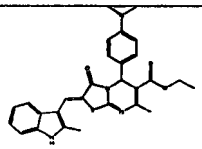
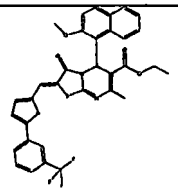
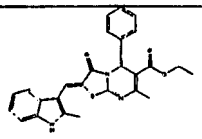
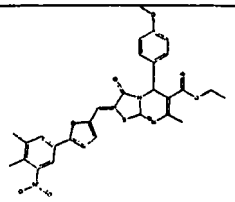
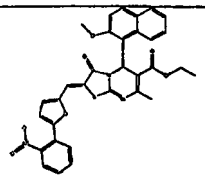
ID	STRUCTURE	MarA Lance	GoxG Lance	Rob Lance	OD	CV	Luc
SD		*	***	*	**	ND	*
SE		**	*	*	**	ND	ND
SF		*	*	**	**	ND	ND
SG		**	*	NT	**	ND	*
SH		**	**	NT	**	ND	ND
SI		**	***	NT	**	ND	ND
SJ		*	***	*	**	**	ND
SK		*	*	*	**	**	ND
SL		*	*	*	**	**	ND

TABLE 5

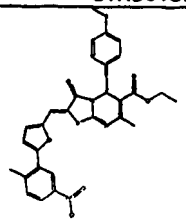
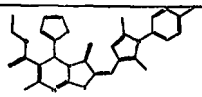
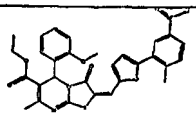
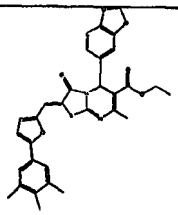
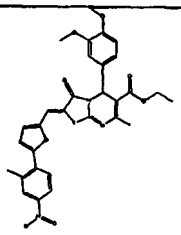
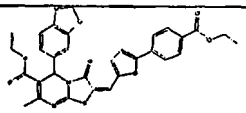
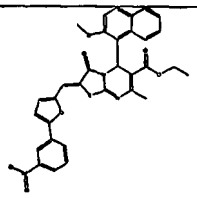
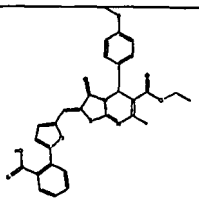
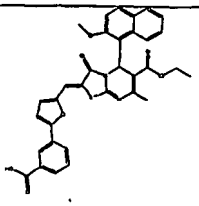
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
SM		*	NT	NT	**	*	**
SN		*	NT	NT	**	*	*
SO		*	NT	NT	**	**	*
SP		NT	**	**	**	ND	**
SQ		NT	**	***	**	ND	**
SR		NT	*	*	**	ND	*
SS		NT	*	*	**	ND	ND
ST		NT	**	*	**	ND	**
SU		NT	*	*	**	ND	**

TABLE 5

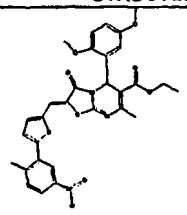
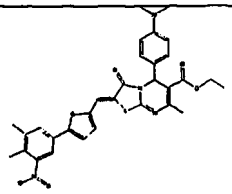
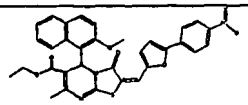
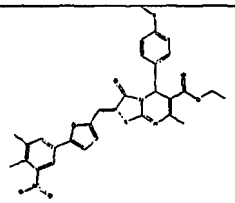
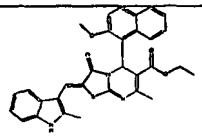
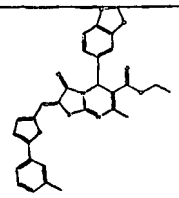
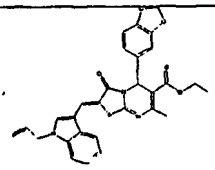
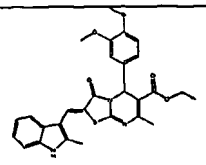
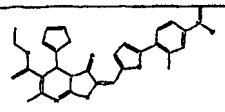
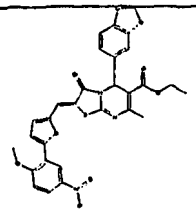
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
SV		**	**	*	**	**	**
SW		*	*	**	**	**	*
SX		*	NT	NT	**	ND	*
SY		***	*	*	**	ND	ND
SZ		*	*	*	**	ND	ND
TA		*	*	*	**	ND	*
TB		*	*	*	**	ND	**
TC		*	*	*	**	ND	ND
TD		*	*	*	**	ND	*
TE		*	*	*	**	ND	**

TABLE 5

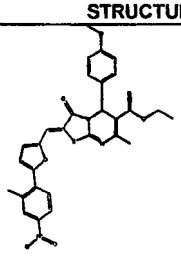
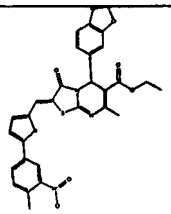
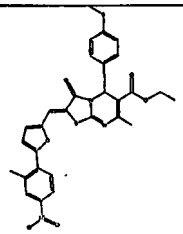
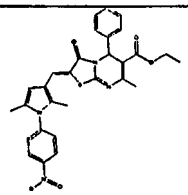
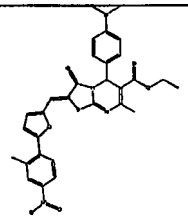
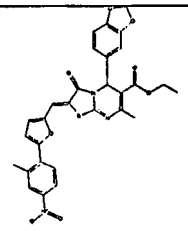
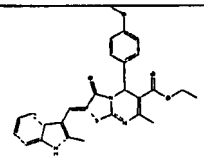
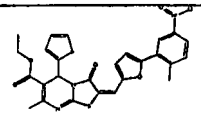
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Loc
TF		*	*	*	**	ND	**
TG		***	*	**	**	ND	**
TH		*	*	*	**	ND	**
TI		*	*	**	**	ND	*
TJ		*	*	*	**	ND	**
TK		*	**	***	**	ND	ND
TL		*	*	*	**	ND	ND
TM		***	***	*	**	ND	*

TABLE 5

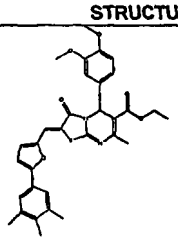
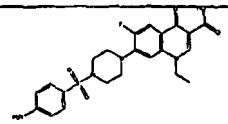
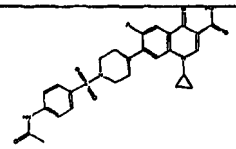
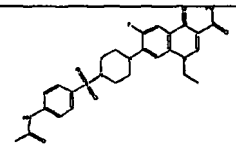
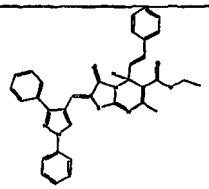
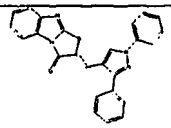
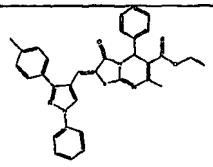
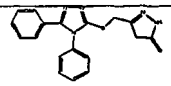
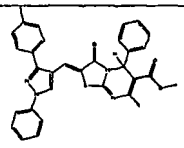
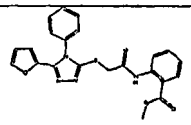
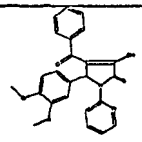
ID	STRUCTURE	MazA Lance	BoxS Lance	Rob Lance	oo	cv	Luc
TN		*	*	*	**	ND	ND
TO		*	*	***	**	ND	*
TP		*	*	**	**	ND	ND
TQ		*	**	***	**	ND	*
TR		*	**	**	**	ND	ND
TS		*	**	*	**	ND	ND
TT		*	***	***	**	ND	ND
TU		**	*	*	**	ND	*
TV		*	*	*	**	ND	*
TW		*	*	*	**	ND	*
TX		**	*	*	**	ND	*

TABLE 5

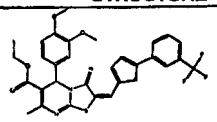
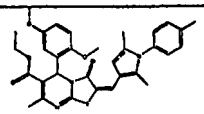
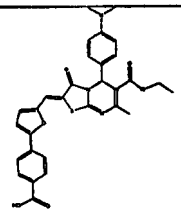
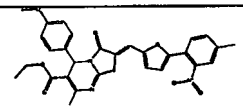
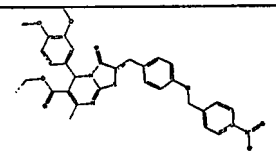
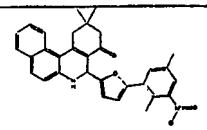
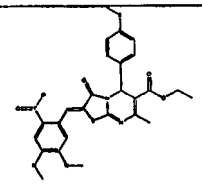
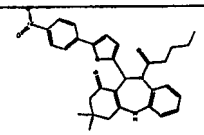
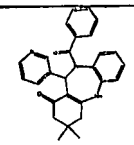
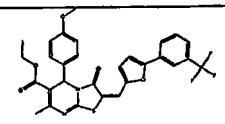
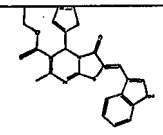
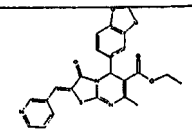
ID	STRUCTURE	MarA Lance	BoxS Lance	Rob Lance	OD	CV	Luc
TY		*	*	*	**	ND	ND
TZ		*	*	*	**	ND	ND
UA		*	*	*	**	ND	ND
UB		**	*	*	**	*	*
UC		**	***	***	**	**	**
UD		**	***	***	**	*	**
UE		*	*	**	**	ND	**
UF		**	**	*	**	ND	**
UG		**	**	**	**	ND	**
UH		***	***	***	**	ND	**
UI		**	*	**	**	ND	*
UJ		**	*	*	**	ND	ND

TABLE 5

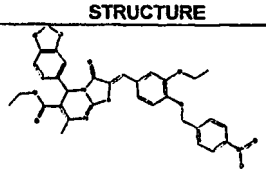
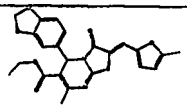
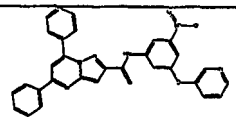
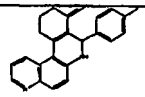
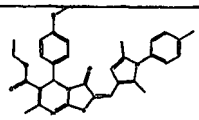
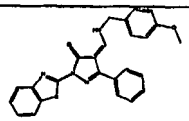
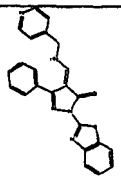
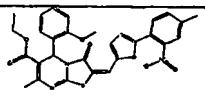
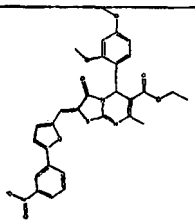
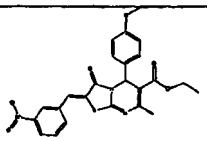
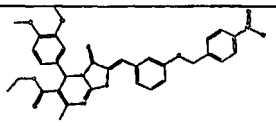
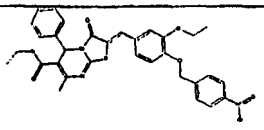
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lux
UK		***	***	*	**	ND	*
UL		**	**	**	**	ND	*
UM		**	*	***	**	ND	**
UN		***	*	**	**	ND	ND
UO		***	*	*	**	ND	**
UP		***	*	*	**	ND	ND
UQ		**	*	*	**	ND	**
UR		*	*	*	**	ND	**
US		**	*	*	**	ND	**
UT		*	*	*	**	ND	*
UU		*	**	*	**	ND	ND
UV		***	***	NT	**	ND	**

TABLE 5

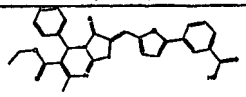
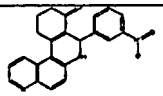
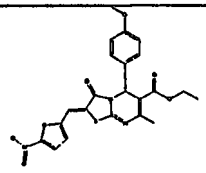
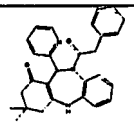
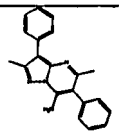
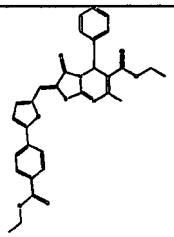
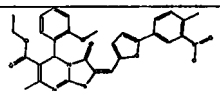
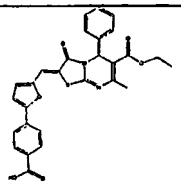
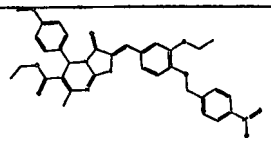
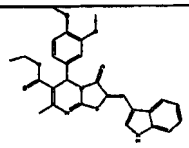
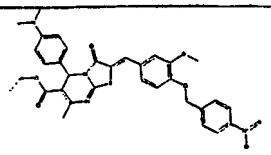
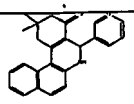
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
UW		***	**	NT	**	ND	**
UX		***	***	NT	**	ND	*
UY		***	*	NT	**	ND	**
UZ		***	**	NT	**	ND	**
VA		*	*	*	**	**	**
VB		*	*	*	**	**	**
VC		*	*	*	**	**	**
VD		*	*	*	**	**	**
VE		*	NT	NT	**	**	*
VF		*	NT	NT	**	**	*
VG		*	NT	NT	**	*	ND
VH		**	**	***	**	**	**

TABLE 5

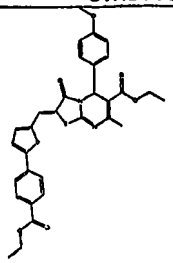
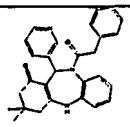
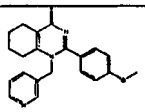
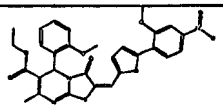
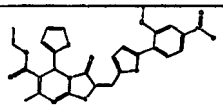
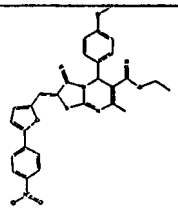
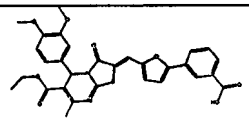
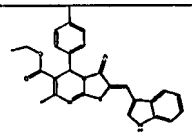
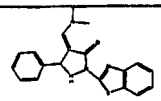
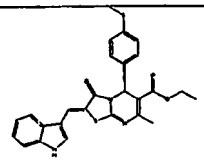
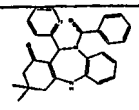
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
VI		*	*	*	**	*	*
VJ		*	**	*	**	**	**
VK		*	**	**	**	**	**
VL		NT	**	***	**	ND	*
VM		NT	*	*	**	ND	*
VN		NT	**	*	**	ND	*
VO		*	***	*	**	**	*
VP		***	NT	NT	**	ND	ND
VQ		**	NT	NT	**	ND	ND
VR		*	NT	NT	**	ND	ND
VS		*	*	**	**	ND	**

TABLE 5

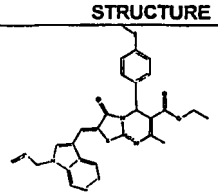
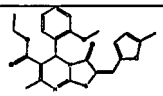
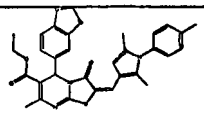
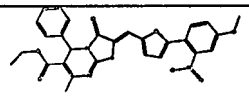
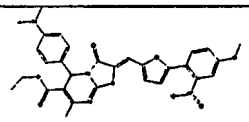
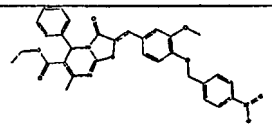
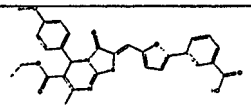
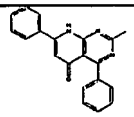
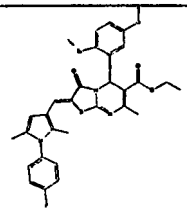
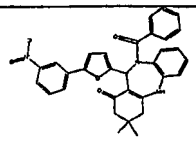
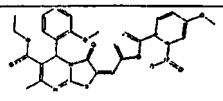
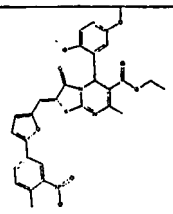
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
VT		*	*	***	**	ND	*
VU		***	***	***	**	ND	**
VV		*	*	***	**	ND	*
VW		*	*	*	**	ND	**
VX		*	*	**	**	ND	**
VY		*	*	*	**	ND	ND
VZ		*	*	*	**	ND	**
WA		**	*	*	**	ND	**
WB		**	***	*	**	ND	**
WC		*	**	**	**	ND	**
WD		***	*	*	**	ND	**
WE		**	**	*	**	ND	ND

TABLE 5

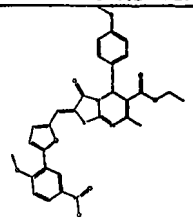
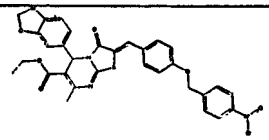
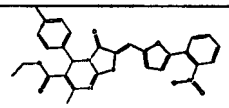
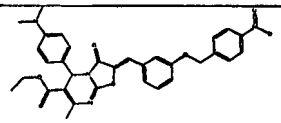
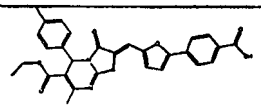
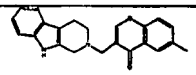
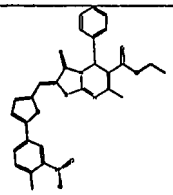
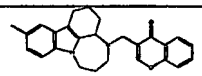
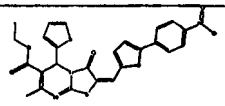
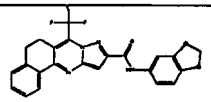
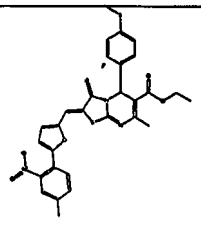
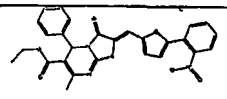
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
WF		**	***	**	**	ND	**
WG		*	*	*	**	ND	ND
WH		**	*	**	**	ND	**
WI		*	**	***	**	ND	**
WJ		*	*	*	**	ND	*
WK		*	*	*	**	ND	*
WL		*	*	*	**	ND	**
WM		*	*	**	**	ND	*
WN		*	*	*	**	ND	*
WO		*	**	***	**	ND	**
WP		*	*	**	**	ND	*
WQ		*	***	***	**	ND	ND

TABLE 5

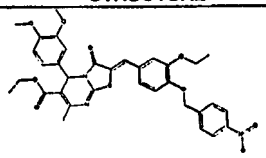
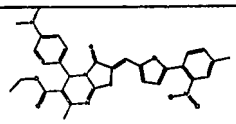
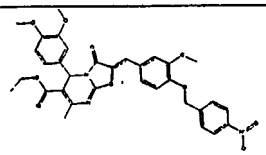
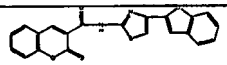
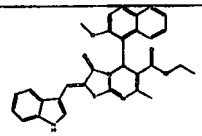
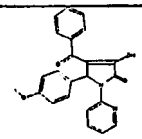
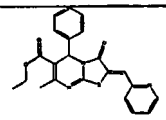
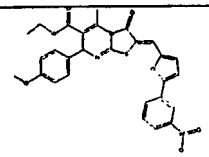
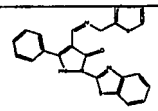
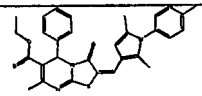
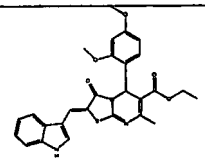
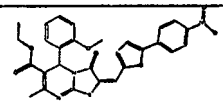
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
WR		•	•	•	••	ND	ND
WS		•	•	•	••	ND	ND
WT		••	•	•	••	ND	••
WU		•	•	•	••	ND	•
WV		•	•	•	••	ND	••
WW		•	•	•	••	ND	•
WX		••	•	•	••	ND	•
WY		•	•	•	••	ND	••
WZ		•	•	•	••	ND	ND
XA		•	•	•	••	ND	ND
XB		•	•	•	••	ND	•
XC		•	•	•	••	ND	ND

TABLE 5

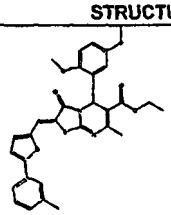
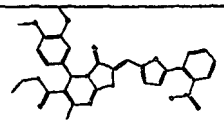
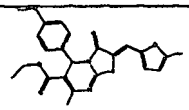
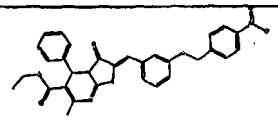
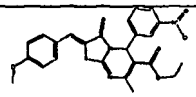
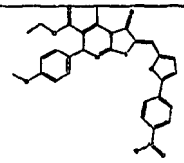
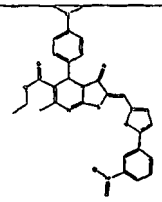
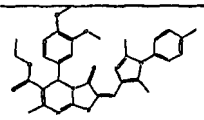
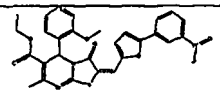
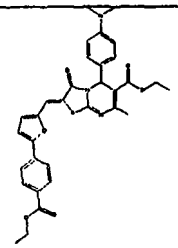
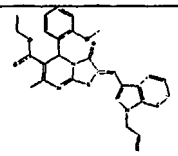
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
XD		***	*	*	**	ND	ND
XE		*	*	**	**	ND	ND
XF		***	***	**	**	*	ND
XG		*	**	*	**	*	*
XH		*	**	*	**	**	**
XI		**	**	***	**	*	**
XG		*	**	*	**	**	*
XH		*	*	*	**	*	ND
XI		*	*	*	**	**	*
XJ		*	**	***	**	**	**
XK		*	*	*	**	*	*

TABLE 5

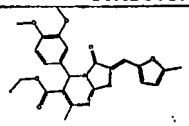
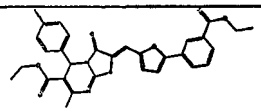
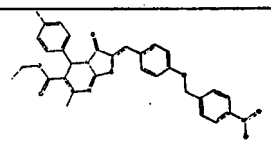
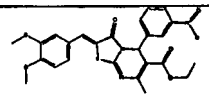
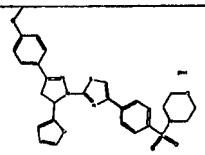
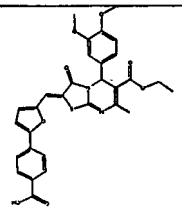
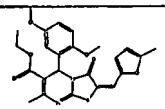
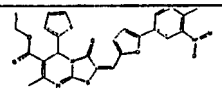
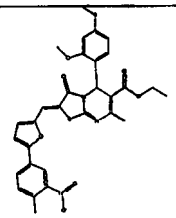
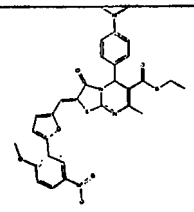
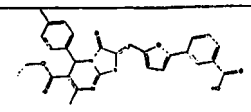
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
XL		**	**	**	**	ND	*
XM		***	*	*	**	ND	**
XN		**	*	*	**	ND	ND
XO		***	***	*	**	ND	**
XP		**	*	**	**	ND	ND
XQ		**	*	**	**	ND	*
XR		***	*	*	**	ND	**
XS		**	*	*	**	ND	**
XT		**	*	*	**	ND	**
XU		*	*	*	**	ND	*
XV		**	*	*	**	ND	*

TABLE 5

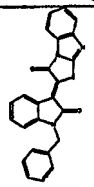
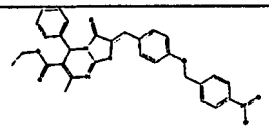
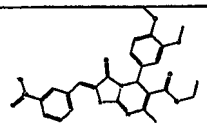
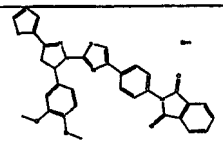
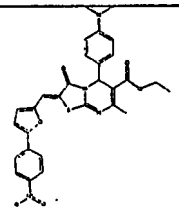
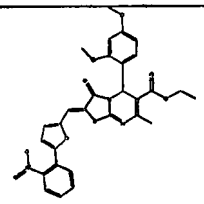
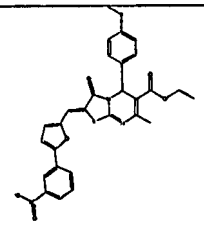
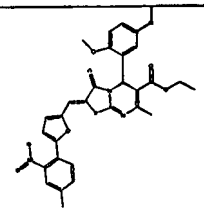
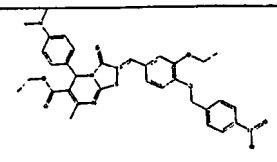
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lux
XW		**	*	**	**	ND	ND
XY		*	***	*	**	ND	**
XZ		***	***	NT	**	ND	**
YA		**	*	*	**	**	*
YB		*	*	***	**	**	**
YC		*	*	**	**	**	**
YD		*	NT	NT	**	**	**
YE		*	NT	NT	**	**	*
YF		***	***	***	**	**	**

TABLE 5

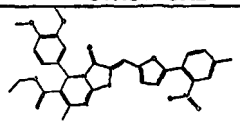
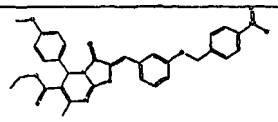
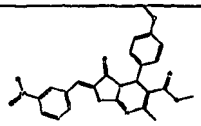
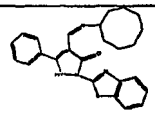
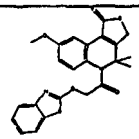
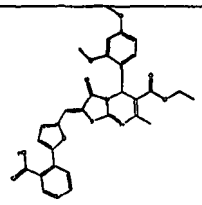
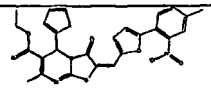
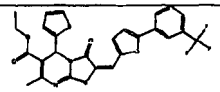
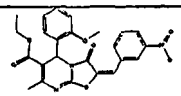
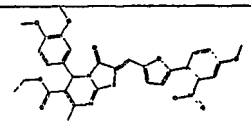
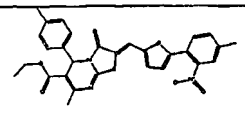
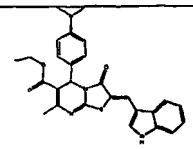
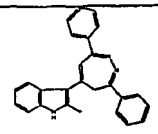
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
YG		*	**	***	**	**	*
YH		***	***	*	**	**	**
YI		***	***	*	**	**	**
YJ		NT	**	*	**	ND	**
YK		NT	***	*	**	ND	*
YL		NT	**	*	**	ND	**
YM		NT	**	*	**	ND	*
YN		NT	**	**	**	ND	**
YO		*	*	*	**	**	ND
YP		*	*	*	**	**	**
YQ		*	**	*	**	**	**
YR		**	NT	NT	**	ND	ND
YS		**	*	***	**	ND	*

TABLE 5

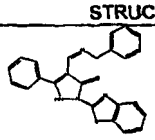
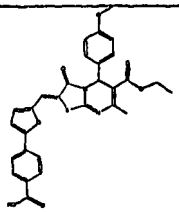
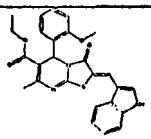
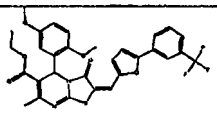
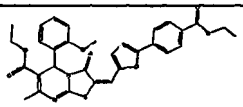
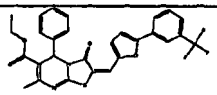
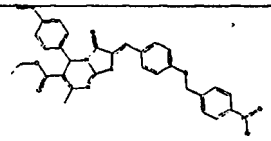
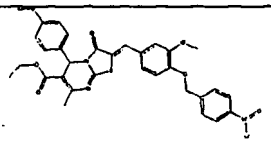
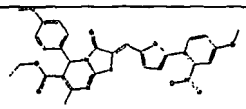
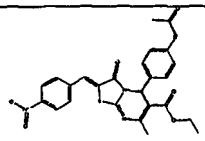
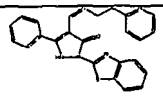
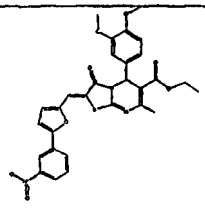
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	CD	CV	Loc
YT		***	*	*	**	ND	*
YU		**	*	**	**	ND	**
YV		*	*	*	**	ND	**
YW		***	***	***	**	ND	**
YX		*	*	**	**	ND	**
YY		*	*	*	**	ND	**
YZ		**	*	*	**	ND	*
ZA		*	*	*	**	ND	**
ZB		*	*	*	**	ND	ND
ZC		***	***	***	**	ND	**
ZD		***	***	***	**	ND	ND
ZE		**	*	***	**	ND	ND

TABLE 5

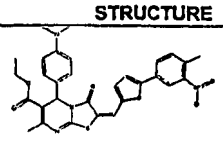
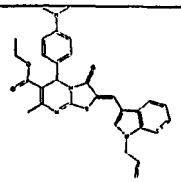
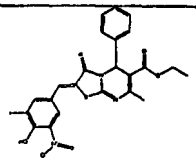
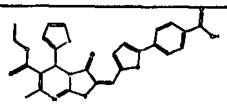
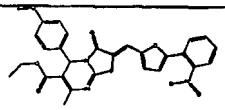
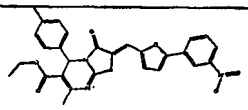
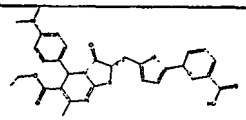
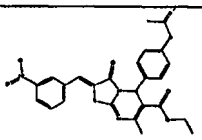
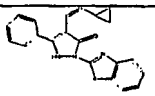
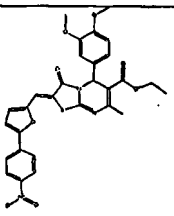
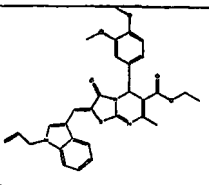
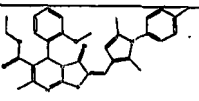
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ZF		**	*	*	**	ND	*
ZG		**	*	***	**	ND	ND
ZH		*	**	***	**	ND	*
ZI		*	*	*	**	ND	ND
ZJ		*	*	*	**	ND	*
ZK		*	*	*	**	*	ND
ZL		*	*	*	**	*	ND
ZM		*	**	**	**	*	ND
ZN		*	**	*	**	*	ND
ZO		*	*	*	**	*	**
ZP		*	**	***	**	**	**
ZQ		**	***	***	**	*	ND

TABLE 5

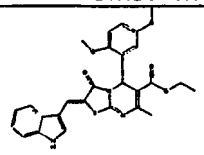
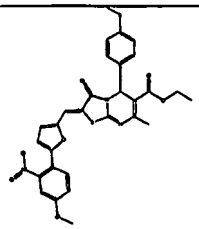
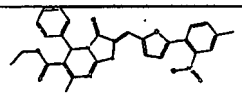
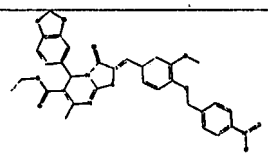
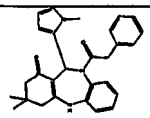
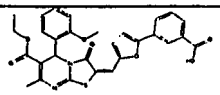
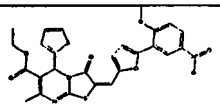
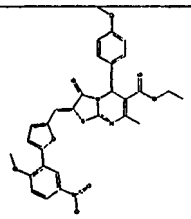
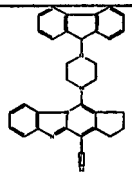
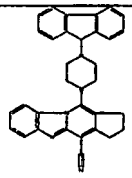
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ZR		***	***	*	**	**	**
ZS		*	***	*	**	ND	**
ZT		**	*	**	**	ND	**
ZU		*	*	*	**	ND	**
AAA		*	*	*	**	ND	*
AAB		***	***	*	**	ND	ND
AAC		**	*	*	**	ND	ND
AAD		***	*	*	**	ND	*
AAE		**	*	*	**	ND	**
AAF		***	*	*	**	ND	*

TABLE 5

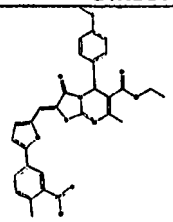
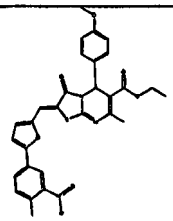
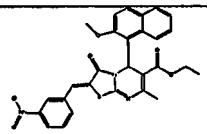
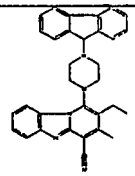
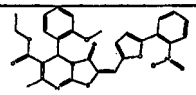
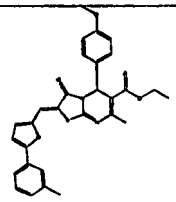
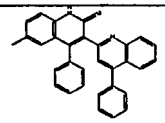
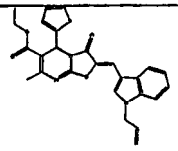
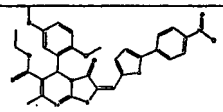
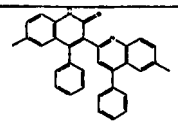
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AAG		***	*	*	**	ND	**
AAH		**	*	*	**	ND	**
AAI		**	*	***	**	ND	ND
AAJ		***	**	NT	**	ND	ND
AAK		***	*	NT	**	ND	*
AAL		*	*	*	**	ND	ND
AAM		**	***	**	**	**	ND
AAN		*	*	*	**	**	ND
AAO		*	*	*	**	*	*
AAP		*	*	*	**	*	*

TABLE 5

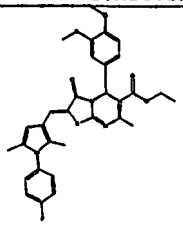
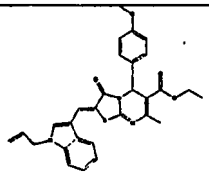
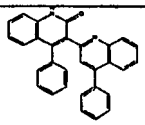
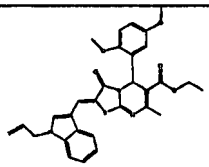
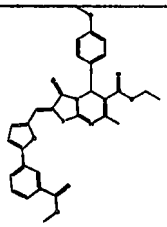
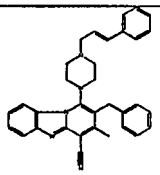
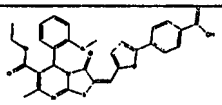
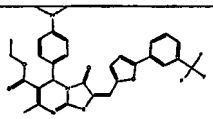
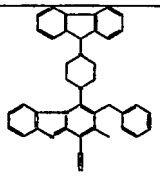
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	CD	CV	Loc
AAQ		**	*	*	**	**	**
AAR		*	*	*	**	**	**
AAS		*	**	***	**	**	*
AAT		*	***	**	**	**	*
AAU		NT	*	*	**	ND	*
AAV		NT	*	***	**	ND	**
AAW		NT	***	**	**	ND	ND
AAX		NT	***	**	**	ND	**
AAZ		NT	*	*	**	ND	ND

TABLE 5

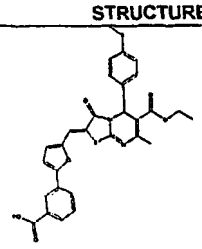
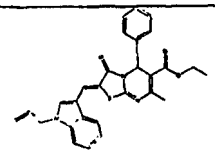
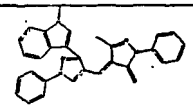
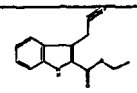
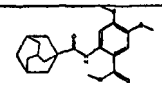
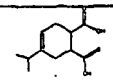
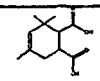
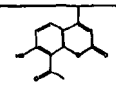
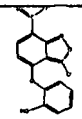
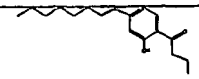
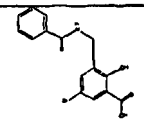
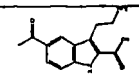
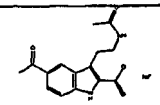
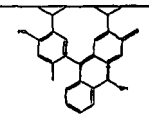
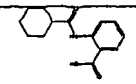
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ABA		NT	**	*	**	ND	**
ABB		NT	***	***	**	ND	**
ABC		***	***	*	**	**	**
ABD		*	*	*	**	*	ND
ABE		*	**	*	**	*	**
ABF		*	*	*	**	**	ND
ABG		*	**	***	**	*	ND
ABH		*	*	*	**	**	*
ABI		**	NT	NT	**	ND	**
ABJ		*	*	***	**	ND	*
ABK		*	*	**	**	ND	**
ABL		*	*	*	**	ND	**
ABN		*	*	*	**	ND	**
ABO		*	*	***	**	ND	**
ABP		***	*	*	**	ND	**

TABLE 5

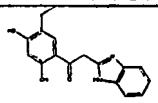
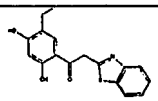
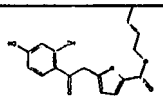
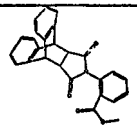
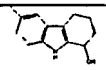
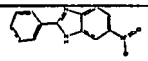
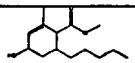
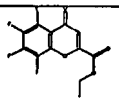
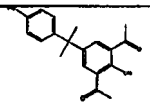
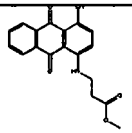
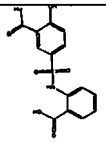
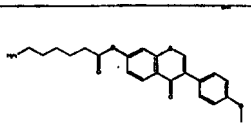
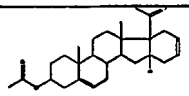
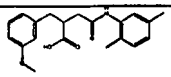
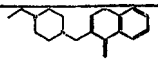
ID	STRUCTURE	MerA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ABQ		*	*	*	**	ND	**
ABR		*	*	*	**	ND	**
ABS		*	*	*	**	ND	**
ABT		*	*	*	**	ND	*
ABU		*	*	*	**	ND	**
ABV		*	**	***	**	ND	**
ABW		*	*	*	**	ND	**
ABX		***	***	**	**	ND	**
ABY		*	*	*	**	ND	**
ABZ		*	**	***	**	ND	*
ACA		*	**	***	**	ND	*
ACB		*	*	*	**	ND	ND
ACC		*	*	*	**	ND	ND
ACD		*	*	*	**	ND	*
ACE		*	*	*	**	ND	ND
ACF		*	*	*	**	ND	*

TABLE 5

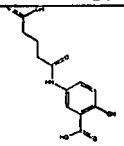
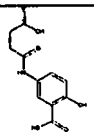
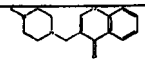
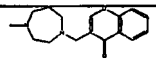
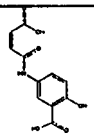
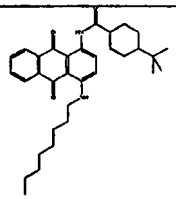
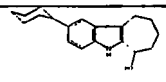
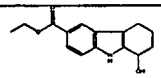
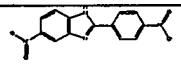
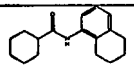
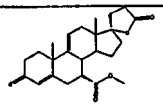
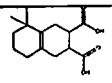
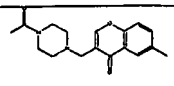
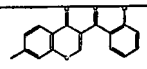
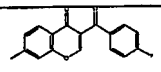
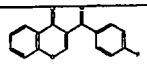
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ACG		*	*	*	**	ND	**
ACH		**	*	*	**	ND	ND
ACI		*	**	***	**	*	**
ACJ		*	**	*	**	*	ND
ACK		*	*	*	**	*	ND
ACL		*	*	*	**	*	ND
ACM		***	***	***	**	ND	**
CAN		***	***	**	**	*	**
ACO		***	***	*	**	*	**
ACP		***	***	**	**	**	**
ACQ		**	***	**	**	**	**
ACR		***	***	*	**	**	**
ACS		*	*	*	**	ND	**
ACT		**	*	**	**	ND	**
ACU		**	*	***	**	ND	**
ACV		*	**	*	**	ND	**
ACW		**	**	**	**	ND	*

TABLE 5

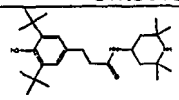
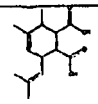
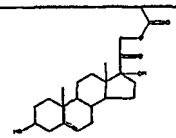
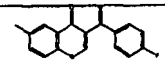
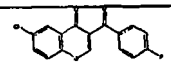
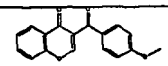
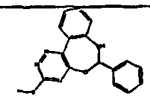
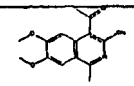
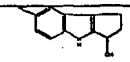
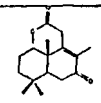
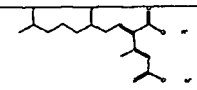
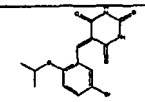
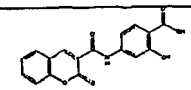
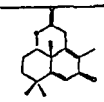
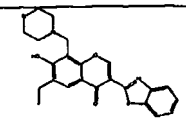
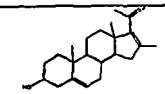
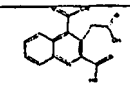
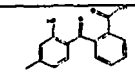
ID	STRUCTURE	NatA Lance	SerS Lance	RodP Lance	CD	CV	Luc
ACX		**	**	**	**	ND	*
ACY		**	*	**	**	ND	ND
ACZ		**	*	*	**	ND	*
ADA		***	***	***	**	ND	ND
ADB		**	*	***	**	ND	ND
ADC		***	**	***	**	ND	*
ADD		***	***	***	**	ND	**
ADE		**	*	*	**	ND	*
ADF		**	*	*	**	ND	**
ADG		*	*	*	**	ND	**
ADH		**	*	*	**	ND	ND
ADI		**	*	*	**	ND	*
ADJ		**	*	*	**	ND	**
ADK		**	*	*	**	ND	**
ADL		*	*	*	**	ND	**
ADM		**	*	*	**	ND	*
ADN		**	*	*	**	ND	*
ADO		**	*	*	**	ND	ND

TABLE 5

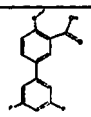
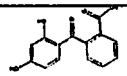
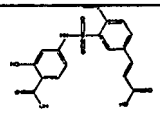
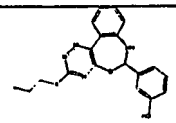
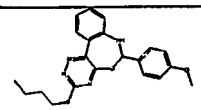
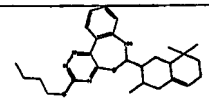
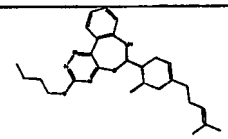
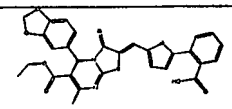
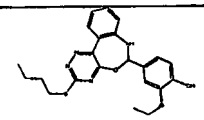
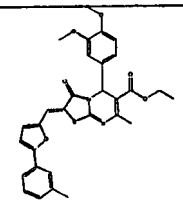
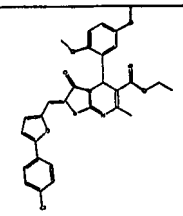
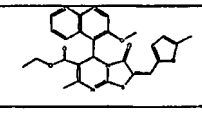
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ADP		**	**	***	**	ND	*
ADQ		**	***	*	**	ND	ND
ADR		***	**	NT	**	ND	**
ADS		*	**	**	*	**	**
ADT		*	*	*	*	**	**
ADU		NT	*	**	*	*	**
ADV		NT	**	*	*	*	ND
ADW		NT	*	*	*	*	ND
ADX		***	***	*	*	*	*
ADY		*	*	*	*	**	**
ADZ		*	**	**	*	**	**
AEA		*	*	*	*	*	**

TABLE 5

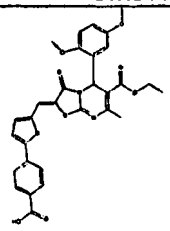
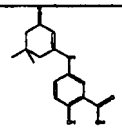
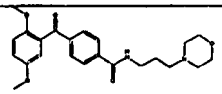
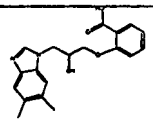
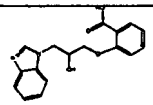
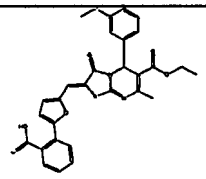
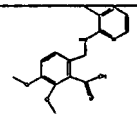
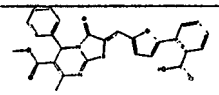
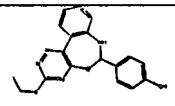
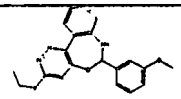
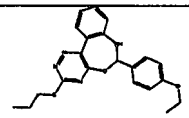
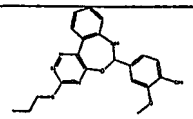
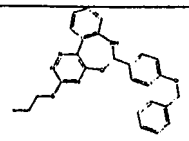
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AEB		*	*	***	*	*	**
AEC		*	*	**	*	*	**
AED		***	*	*	*	*	**
AEF		*	*	*	*	*	ND
EG		*	**	***	*	*	**
AEH		*	*	*	*	*	**
AEI		*	*	*	*	*	*
AEJ		*	**	***	*	*	*
AEK		*	**	**	*	*	ND
AEL		*	*	*	*	*	ND
AEM		*	*	**	*	*	**
EN		*	*	*	*	**	ND
AEO		*	*	*	*	*	**

TABLE 5

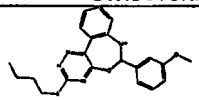
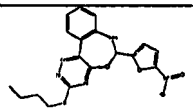
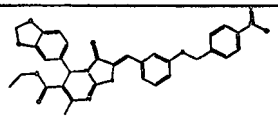
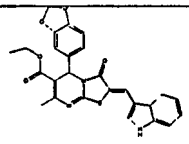
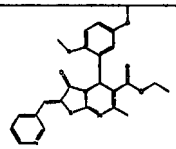
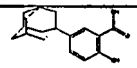
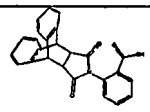
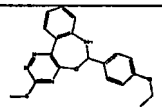
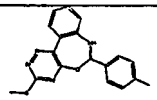
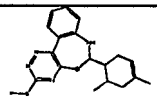
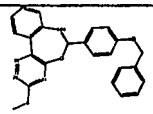
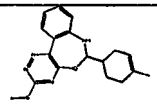
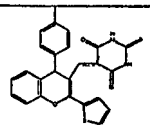
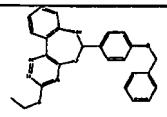
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AEQ		+	+	+	+	++	++
ER		+	+++	+++	+	++	++
AES		+++	+	+	+	+	++
AET		+++	+++	+	+	+	++
EU		+++	+++	+	+	++	++
AEV		++	++	+	+	+	+
AEW		++	+	+	+	+	+
EX		++	++	+	+	+	ND
AEY		++	++	++	+	+	++
AEZ		++	+	+	+	+	++
AFA		++	+	++	+	+	+
FB		++	+	++	+	+	++
AFC		++	+	+	+	+	++
AFD		+	+	+	+	+	++

TABLE 5

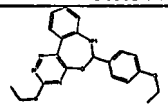
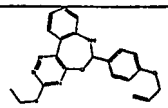
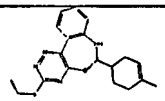
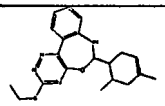
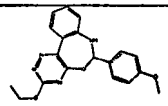
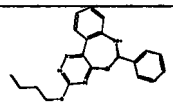
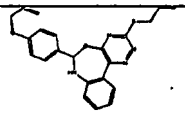
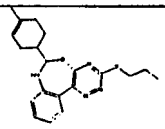
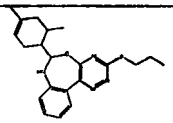
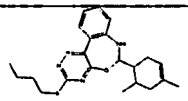
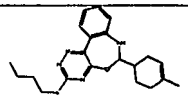
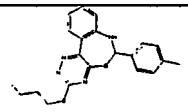
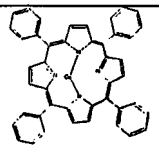
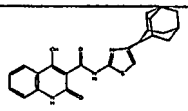
ID	STRUCTURE	MarA Lance	Sox3 Lance	Rob Lance	OD	CV	Luc
AFE		*	*	*	*	*	**
AFG		**	*	*	*	*	**
AFH		*	*	*	*	*	ND
AFI		**	**	*	*	*	**
AFJ		*	**	**	*	**	*
AFK		*	*	**	*	**	**
AFL		*	*	*	*	**	*
AFM		*	*	*	*	**	**
AFN		*	NT	NT	*	**	*
AFO		*	*	***	*	**	ND
AFP		*	*	*	*	**	**
AFQ		*	*	*	*	**	**
AFR		NT	***	***	*	*	*
AFS		NT	**	**	*	*	*

TABLE 5

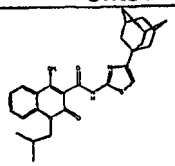
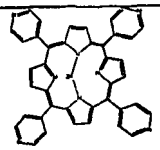
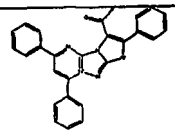
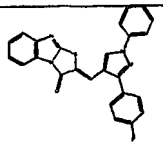
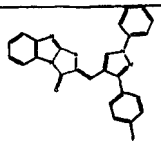
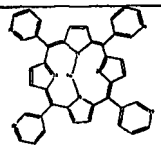
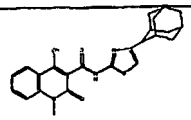
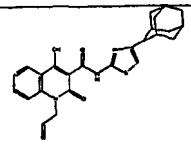
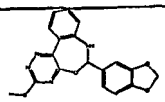
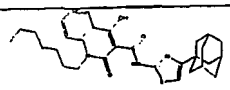
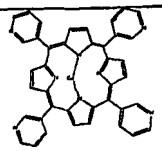
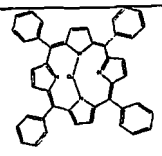
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AFT		NT	*	*	*	*	*
AFU		*	NT	NT	*	*	ND
AFV		*	*	***	*	*	**
AFW		*	*	*	*	*	*
AFX		*	*	*	*	*	**
AFY		***	***	*	*	*	ND
AFZ		*	*	***	*	*	**
AGA		*	*	*	*	*	ND
AGB		*	**	***	*	*	*
AGC		*	**	***	*	*	*
AGD		*	*	*	*	*	ND
AGE		**	**	***	*	*	**

TABLE 5

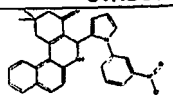
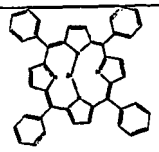
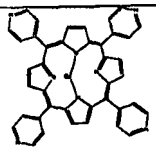
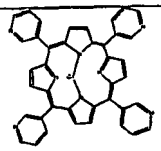
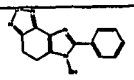
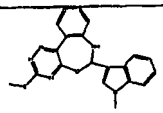
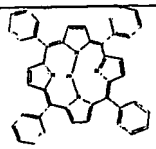
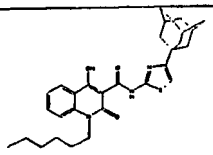
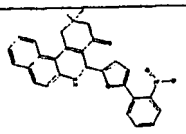
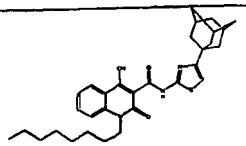
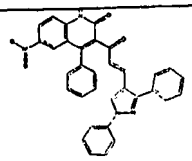
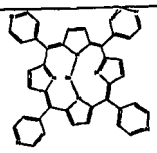
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lac
AGF		•	•	•	•	•	••
AGG		•	••	••	•	•	••
AGH		•	•	••	•	•	••
AGI		•	•	•	•	•	••
AGJ		•	•	•	•	•	••
AGK		•	••	•	•	••	••
AGL		•••	•••	•	•	•	••
AGM		•••	•••	••	•	•	••
AGN		••	•	••	•	•	•
AGO		••	•	•	•	•	•
AGP		••	••	••	•	•	••
AGQ		••	•	•	•	•	••

TABLE 5

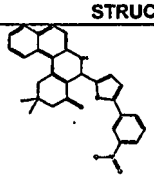
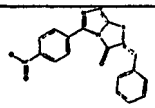
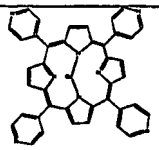
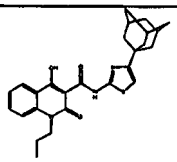
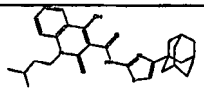
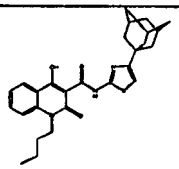
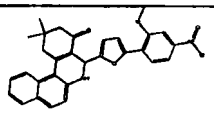
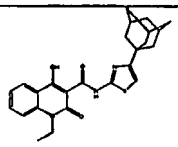
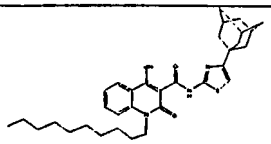
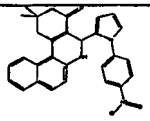
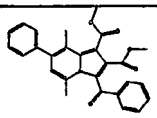
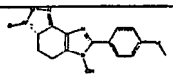
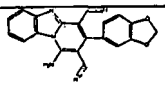
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Lux
AGR		***	*	**	*	*	ND
AGS		***	*	**	*	*	**
AGT		*	*	*	*	*	*
AGU		**	*	*	*	*	ND
AGV		***	***	**	*	*	**
AGW		***	**	NT	*	*	*
AGX		**	**	***	*	**	**
AGY		*	**	*	*	**	*
AGZ		***	***	*	*	**	**
AHA		*	**	*	*	**	**
AHB		NT	**	*	*	*	**
AHC		*	**	*	*	*	*
AHD		*	*	*	*	**	*

TABLE 5

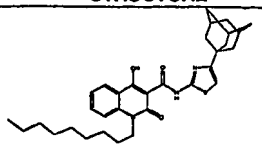
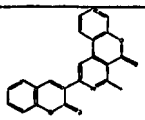
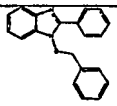
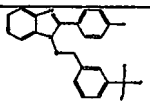
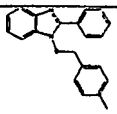
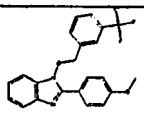
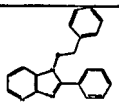
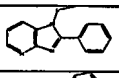
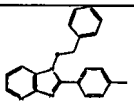
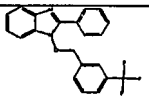
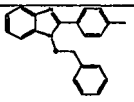
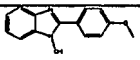
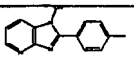
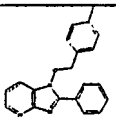
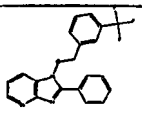
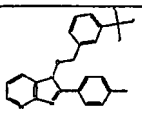
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AHE		*	*	*	*	*	*
AHF		*	*	*	*	*	**
AHG		*	*	*	*	*	**
AHH		*	*	*	*	*	**
AHI		*	*	*	*	*	ND
AHJ		*	*	*	*	*	ND
HK		**	**	***	*	*	**
AHL		*	*	*	*	*	ND
AHM		*	*	*	*	*	ND
AHN		**	**	***	*	*	ND
AHO		*	*	*	*	*	ND
AHP		***	***	**	*	**	**
AHQ		*	*	*	*	*	*
AHR		***	***	***	*	*	*
AHS		**	*	**	*	*	*
AHT		**	*	*	*	*	**

TABLE 5

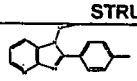
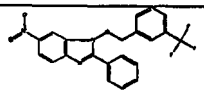
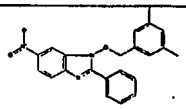
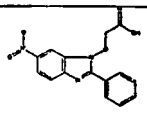
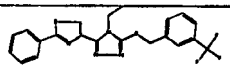
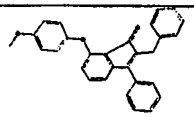
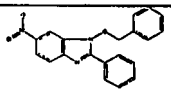
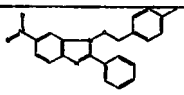
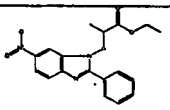
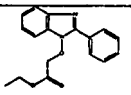
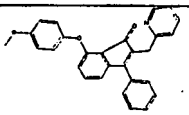
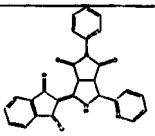
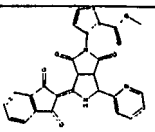
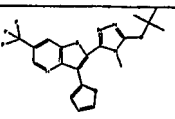
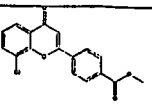
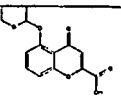
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AHU		**	*	*	*	*	ND
AHV		**	*	*	*	*	**
AHW		**	***	**	*	*	ND
AHX		***	***	NT	*	*	ND
AHY		*	**	***	*	**	**
AHZ		*	NT	NT	*	**	ND
AIA		**	**	*	*	**	ND
AIB		*	*	*	*	*	**
AIC		***	***	***	*	*	**
AID		*	*	*	*	*	ND
AIE		*	*	*	*	*	*
AIF		*	*	*	*	*	ND
AIG		**	*	*	*	*	ND
AIH		*	*	*	*	*	**
AII		*	**	*	*	*	ND
AIJ		***	***	*	*	**	ND

TABLE 5

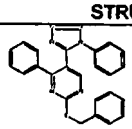
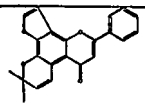
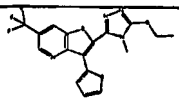
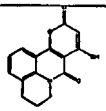
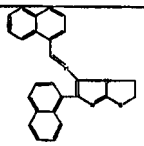
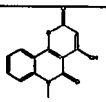
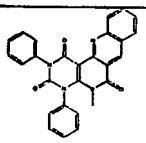
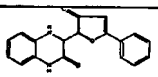
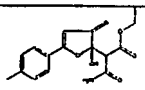
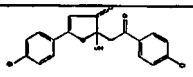
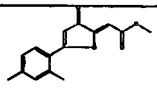
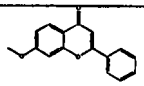
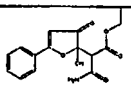
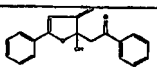
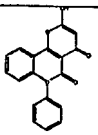
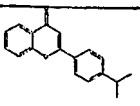
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AIK		**	**	*	*	*	ND
AIL		***	**	***	*	*	*
AIM		*	*	*	*	*	**
AIN		*	*	*	*	*	**
AIO		*	**	**	*	*	**
AIP		**	*	*	*	*	ND
AIQ		**	**	**	*	*	ND
AIR		*	*	*	*	**	**
AIS		*	*	*	*	**	**
AIT		*	*	*	*	**	**
AIU		*	NT	NT	*	*	*
AIV		***	***	**	*	**	**
AIW		NT	***	*	*	*	**
AIX		*	*	*	*	*	**
AIY		*	*	*	*	*	**
AIZ		*	*	*	*	*	ND

TABLE 5

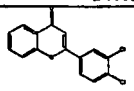
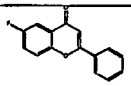
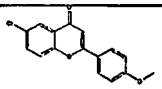
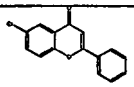
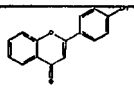
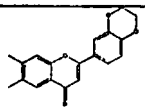
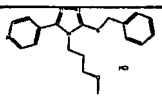
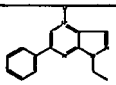
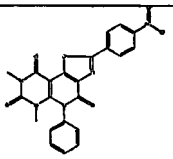
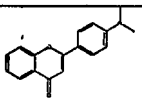
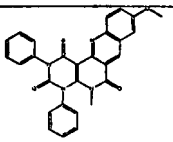
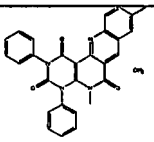
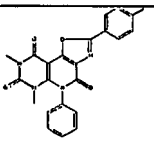
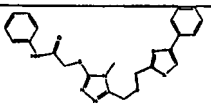
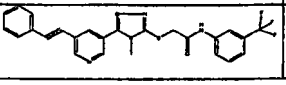
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AJA		*	*	*	*	*	ND
AJB		*	*	*	*	*	ND
AJD		**	*	*	*	*	*
AJE		*	**	*	*	*	ND
AJF		*	*	*	*	*	*
AJG		***	***	***	*	*	**
AJH		***	*	*	*	*	**
AJI		**	*	**	*	*	*
AJJ		**	**	*	*	*	*
AJK		*	*	*	*	*	ND
AJL		**	*	*	*	*	*
AJM		**	**	***	*	*	ND
AJN		***	**	NT	*	*	**
AJO		NT	*	NT	*	*	**
AJP		***	*	NT	*	*	*

TABLE 5

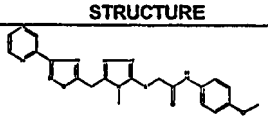
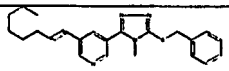
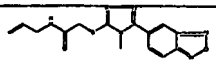
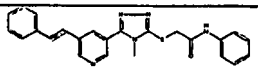
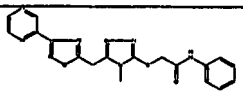
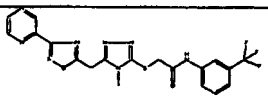
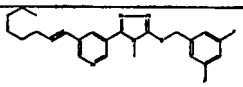
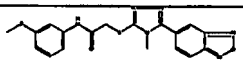
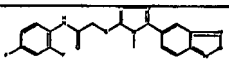
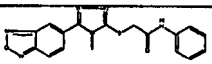
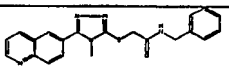
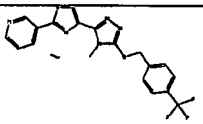
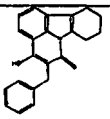
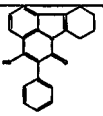
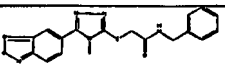
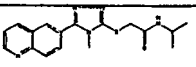
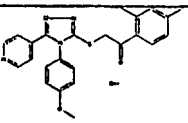
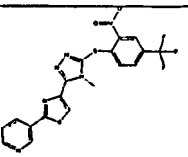
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AJQ		*	*	*	*	*	*
AJR		*	**	***	*	*	**
AJZ		*	NT	NT	*	*	**
AKA		*	*	***	*	*	**
AKB		**	**	***	*	*	**
AKC		*	**	***	*	*	ND
AKD		*	*	*	*	*	ND
AKE		*	**	***	*	*	**
AKF		*	**	***	*	*	*
AKG		*	*	**	*	*	ND
AKH		***	***	***	*	*	**
AKI		*	**	***	*	*	**
AKJ		NT	**	*	*	*	ND
AKK		NT	**	**	*	*	**
AKL		*	**	*	*	*	*
AKM		*	***	***	*	**	**
AKN		**	**	*	*	**	*
AKO		*	*	***	*	*	**

TABLE 5

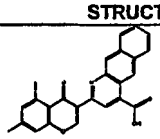
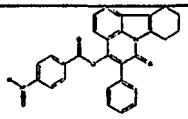
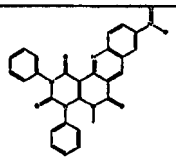
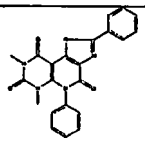
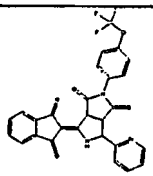
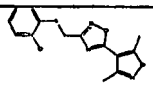
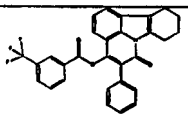
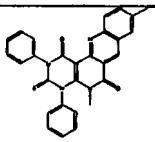
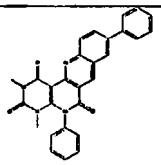
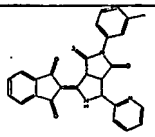
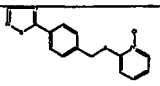
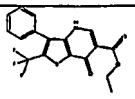
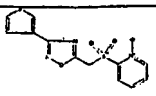
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AKP		NT	**	***	*	*	ND
AKQ		**	**	*	*	*	**
AKR		***	*	*	*	*	*
AKS		***	*	**	*	*	**
AKT		**	***	***	*	*	**
AKU		**	***	***	*	*	**
AKV		**	**	**	*	**	*
AKW		***	***	*	*	**	**
AKX		***	***	***	*	**	**
AKY		*	*	*	*	**	**
AKZ		**	**	**	*	**	**
ALA		*	*	*	*	*	**
ALB		NT	*	*	*	*	ND

TABLE 5

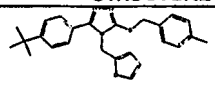
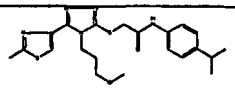
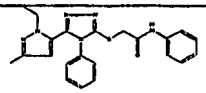
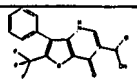
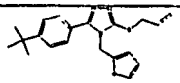
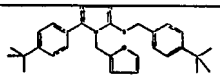
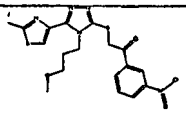
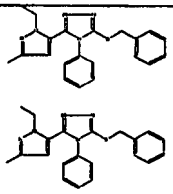
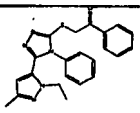
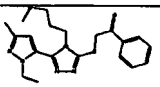
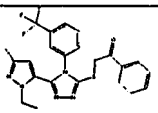
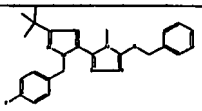
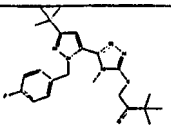
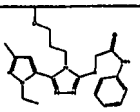
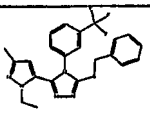
ID	STRUCTURE	MarA Lance	Box3 Lance	Rob Lance	OD	CV	Luc
ALC		**	*	*	*	*	**
ALD		**	NT	NT	*	*	ND
ALE		*	*	*	*	*	**
ALF		**	**	***	*	*	ND
ALG		***	**	***	*	*	*
ALH		*	*	*	*	*	*
ALI		**	*	*	*	*	**
ALJ		**	*	*	*	*	**
ALK		NT	**	***	*	*	**
ALL		*	***	***	*	*	*
ALM		**	*	*	*	*	**
ALN		*	**	*	*	**	**
ALO		*	***	***	*	*	*
ALP		**	*	*	*	*	*
ALQ		***	***	***	*	*	*

TABLE 5

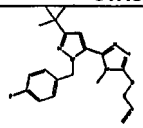
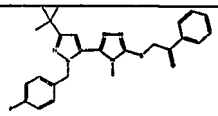
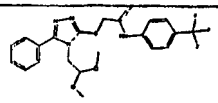
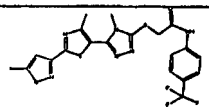
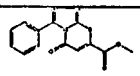
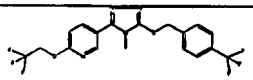
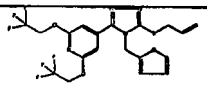
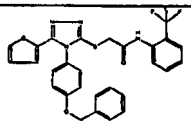
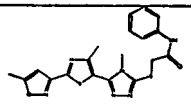
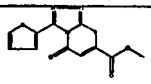
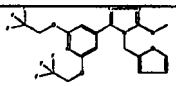
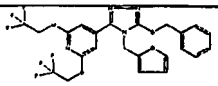
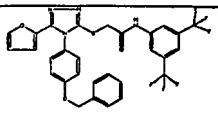
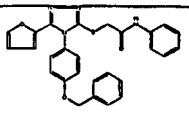
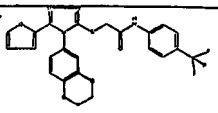
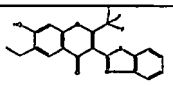
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
ALR		*	**	*	*	*	*
ALS		*	NT	NT	*	*	**
ALT		***	**	NT	*	*	**
ALU		*	**	***	*	**	ND
ALV		**	**	***	*	**	ND
ALW		*	*	*	*	**	ND
ALX		**	***	*	*	**	*
ALY		**	*	*	*	**	**
ALZ		*	*	*	*	**	**
AMA		*	NT	NT	*	**	*
AMB		***	***	**	*	**	**
AMC		*	**	**	*	**	**
AMD		*	*	*	*	**	*
AME		*	*	*	*	**	*
AMF		*	*	*	*	**	**
AMG		***	***	***	*	**	ND

TABLE 5

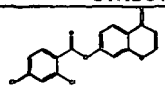
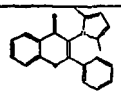
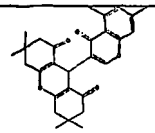
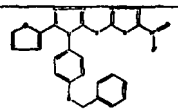
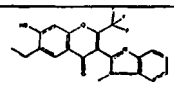
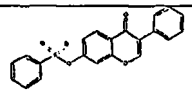
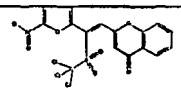
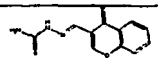
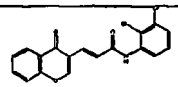
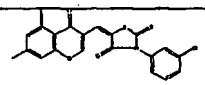
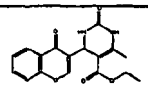
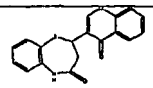
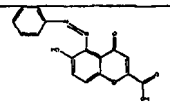
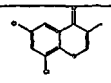
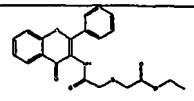
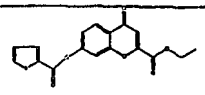
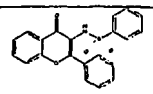
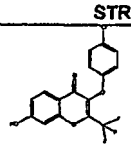
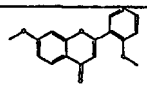
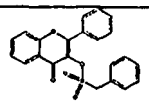
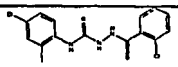
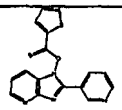
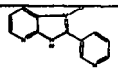
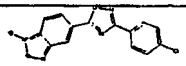
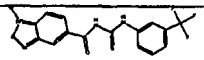
ID	STRUCTURE	MarA Lance	SoxS Lance	Rob Lance	OD	CV	Luc
AMH		*	NT	NT	*	**	**
AMI		*	*	*	*	**	**
AMJ		*	*	***	*	**	ND
AMK		**	***	*	*	**	**
AML		*	*	*	*	**	ND
AMM		**	*	**	*	**	ND
AMN		*	**	*	*	*	*
AMO		*	*	**	*	**	**
AMP		*	*	**	*	**	ND
AMQ		*	**	*	*	**	ND
AMR		*	NT	NT	*	*	*
AMS		**	*	**	*	**	ND
AMT		*	**	**	*	*	**
AMU		***	***	*	*	**	**
AMV		*	*	*	*	*	*
AMX		NT	***	***	*	**	**
AMY		***	*	***	*	**	**
AMZ		*	*	*	*	**	ND

TABLE 5

ID	STRUCTURE	MarA Lance	BoxS Lance	Rob Lance	OD	CV	Luc
ANA		*	*	*	*	**	**
ANB		*	*	**	*	**	**
ANC		*	**	**	*	**	**
AND		NT	***	***	*	**	**
ANE		*	*	*	*	**	**
ANF		***	***	***	*	**	**
ANG		*	*	*	*	**	**
ANH		**	***	NT	*	**	*

CLAIMS

1. A method for reducing antibiotic resistance of a microbial cell, comprising:

contacting said cell with a transcription factor modulating compound, such that the antibiotic resistance of said cell is reduced.

2. The method of claim 1, wherein said transcription factor modulating compound is of the formula (I):

A-E (I)

wherein

A is a polar moiety;

E is a hydrophobic moiety, and pharmaceutically acceptable salts thereof.

3. A method for modulating a transcription, comprising contacting a transcription factor with a transcription factor modulating compound, such that the transcription is modulated, wherein said transcription factor modulating compound is of the formula (I):

A-E (I)

wherein

A is a polar moiety; and

E is a hydrophobic moiety, and pharmaceutically acceptable salts thereof.

4. The method of claim 2 or 3, wherein said polar moiety comprises at least one heterocycle.

5. The method of claim 4, wherein said heterocycle is bicyclic.

6. The method of claim 4, wherein said heterocycle comprises at least one nitrogen atom.

7. The method of claim 6, wherein said heterocycle is selected from the group consisting of benzoimidazole, imidazopyridine, pyridine, pyrollidine, quinoline, triazole, pyrimidine, tetrazole, and porphyin.

8. The method of claim 4, wherein said heterocycle comprises at least one oxygen atom.

9. The method of claim 8, wherein said heterocycle is chromenone.

10. The method of claim 4, wherein said polar moiety is a fused ring moiety.

11. The method of any one of claims 4-10, wherein said heterocycle is substituted.

12. The method of claim 11, wherein the substituent is nitro, alkoxy, aryl, amidyl, ester, thioester, alkyl, araalkyl, halogen, hydroxy, or halogen.

13. The method of claim 12, wherein said substituent is hydroxyl.

14. The method of claim 12, wherein said substituent is halogen.

15. The method of any one of claims 2-3, wherein said hydrophobic moiety comprises at least one alkyl, alkenyl, alkynyl, or aryl moiety.

16. The method of claim 15, wherein said hydrophobic moiety is aryl.

17. The method of claim 16, wherein said hydrophobic moiety is substituted or unsubstituted phenyl.

18. The method of claim 17, wherein said phenyl is substituted with alkyl, alkoxy, halogen, amino, thiol, hydroxy, alkoxy, or nitro.

19. The method of claim 17, wherein said phenyl is *para*-substituted.

20. The method of claim 19, wherein said *para*-substituent is alkyl.

21. The method of claim 20, wherein said alkyl substituent is selected from the group consisting of methyl, ethyl, propyl, butyl, or pentyl.

22. The method of claim 19, wherein said *para*-substituent is hydroxyl.

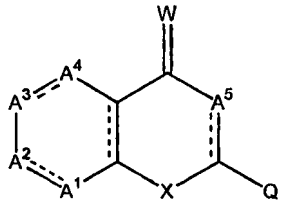
23. The method of claim 19, wherein said *para*-substituent is amino.

24. The method of claim 19, wherein said *para*-substituent is halogen.

25. The method of claim 16, wherein said aryl moiety is heterocyclic.

26. The method of claim 25, wherein said moiety is imidazopyridine, quinolinyl, or pyridinyl.

27. The method of anyone of claims 2-3, wherein said transcription factor modulating compound is of the formula (VII):



(VII)

wherein

W is NH, O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^1, O, or S;

A^2 is C-Z^2, O, or S;

A^3 is C-Z^3, O, or S;

A^4 is C-Z^4, O, or S;

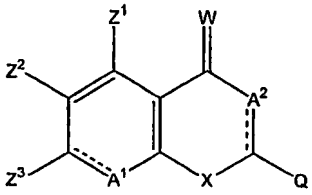
A^5 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, and Z^4 are each independently selected from the group consisting of hydrogen, alkoxy, hydroxy, halogen, and alkyl;

Z^5 is hydrogen, alkoxy, hydroxy, halogen, alkyl, or carbonyl;

Q is hydrogen, alkyl, alkenyl, alkynyl, halogen, hydroxy, aryl, and pharmaceutically acceptable salts thereof.

28. The method of claim 27, wherein said transcription factor modulating compound is of the formula (II):



(II)

wherein

W is O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^4, O, or S;

A^2 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, Z^4 and Z^5 are each independently hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, or cyano;

Z^3 is hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, nitro, cyano, carbonyl, or thiocarbonyl;

Q is an aromatic or heterocyclic moiety, and pharmaceutically acceptable salts thereof.

29. The method of claim 28, wherein W and X are oxygen.

30. The method of claim 28, wherein A^1 is C-Z^4 and Z^4 is hydrogen.

31. The method of claim 28, wherein A^2 is C-Z^5, and Z^5 is hydrogen or hydroxy.

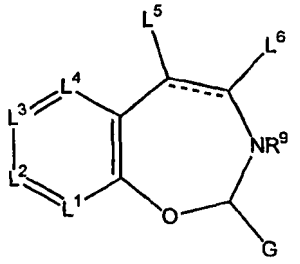
32. The method of claim 28, wherein Z^1 is hydrogen or hydroxy.

33. The method of claim 28, wherein Z^2 is hydrogen or halogen.

34. The method of claim 28, wherein Z^3 is hydrogen, alkoxy or hydroxy.

35. The method of claim 28, wherein Q is substituted phenyl.

36. The method of claim 2 or 3, wherein said transcription factor modulating compound is of the formula (VIII):



(VIII)

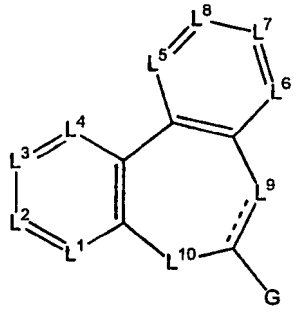
wherein

G is a substituted or unsubstituted aromatic moiety, alkyl, alkenyl, alkynyl, hydrogen;

L^1, L^2, L^3, and L^4 are each independently selected from oxygen, nitrogen, sulfur and or substituted or unsubstituted carbon; and

R^9, L^5 and L^6 are each independently hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, acyl, or aryl, and L_5 and L_6 may optionally be linked with a chain of one to six atoms to form a fused ring, and pharmaceutically acceptable salts thereof.

37. The method of claim 36, wherein said transcription factor modulating compound is of the formula (III):



(III)

wherein

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen; and

L^1, L^2, L^3, L^4, L^5, L^6, L^7, L^8, L^9, and L^{10} are each independently oxygen, substituted or unsubstituted nitrogen, sulfur and or substituted or unsubstituted carbon, and pharmaceutically acceptable salts thereof.

38. The method of claim 36 or 37, wherein R^9 is hydrogen.

39. The method of any one of claims 36-38, wherein G is substituted or unsubstituted phenyl or cyclohexenyl.

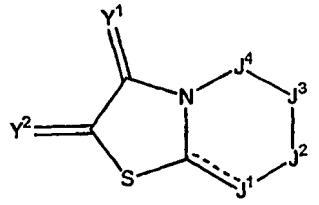
40. The method of any one of claims 36-38, wherein G is heteroaryl.

41. The method of claim 37, wherein L^1, L^2, L^3, and L^4 are each substituted or unsubstituted carbon and L^5, L^6, and L^8 are each nitrogen.

42. The method of claim 37 or 41, wherein L^7 is substituted carbon.

43. The method of claim 42, wherein said substituted carbon is substituted with a thioether moiety.

44. The method of claim 2 or 3, wherein said transcription factor modulating compound is of the formula (X):



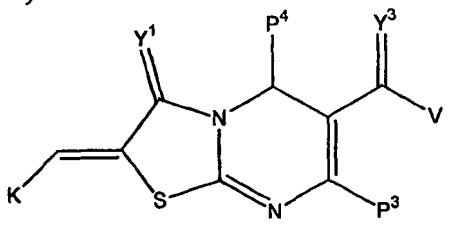
(X)

wherein

Y^1 and Y^2 are each oxygen, sulfur, or substituted or unsubstituted carbon;

J^1, J^2, J^3, and J^4 are each oxygen, nitrogen, or optionally substituted carbon, and pharmaceutically acceptable salts thereof.

45. The method of claim 44, wherein said transcription factor modulating compound is of the formula (IV):



(IV)

wherein

Y^1 and Y^2 are each oxygen or sulfur;

J is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, cyano, nitro, amino, or halogen;

V is substituted or unsubstituted alkyl, alkenyl, alkynyl, alkoxy, alkylamino, or alkylthio;

P and K are each independently substituted or unsubstituted aryl, and pharmaceutically acceptable salts thereof.

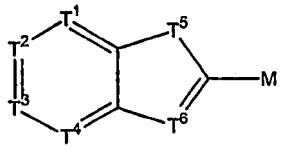
46. The method of claim 45, wherein Y^1 and Y^3 are each oxygen.

47. The method of claim 45 or 46, wherein V is alkoxy and J is lower alkyl.

48. The method of any one of claims 45-47, wherein P is substituted or unsubstituted phenyl.

49. The method of any one of claims 45-48, wherein K is substituted or unsubstituted heteroaryl.

50. The method of claim 2 or 3, wherein said transcription factor modulating compound is of the formula (V):



(V)

wherein

T^1, T^2, T^3, T^4, T^5, and T^6 are each independently substituted or unsubstituted carbon, oxygen, substituted or unsubstituted nitrogen, or sulfur;

M is hydrogen, alkyl, alkenyl, heterocyclic, alkynyl, or aryl, or pharmaceutically acceptable salts thereof.

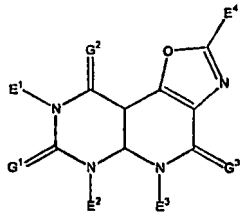
51. The method of claim 50, wherein M is substituted or unsubstituted aryl.

52. The method of claim 50 or 51, wherein T^5 is substituted nitrogen.

53. The method of anyone of claims 50-52, wherein T^1, T^2, T^3 and T^4 are each substituted or unsubstituted carbon.

54. The method of anyone of claims 50-52, wherein one of T^1, T^2, T^3, and T^4 is nitrogen.

55. The method of claims 2 or 3, wherein said transcription factor modulating compound is of the formula (VI):



(VI)

wherein

G^1, G^2, and G^3 are each independently O, S, substituted or unsubstituted nitrogen, or substituted or unsubstituted carbon;

E^1, E^2, and E^3 are each independently hydrogen, alkyl, alkenyl, alkynyl, aryl, aralkyl, or acyl; and

E^4 is alkyl, alkenyl, alkynyl, aryl, halogen, cyano, amino, nitro, or acyl, and pharmaceutically acceptable salts thereof.

56. The method of claim 55, wherein G^1, G^2 and G^3 are each oxygen.

57. The method of claim 55, wherein E^1, E^2, and E^3 are each alkyl.

58. The method of any one of claims 1-57, wherein said transcription factor is a helix-turn-helix protein.

59. The method of any one of claims 1-57, wherein said transcription factor is a transcriptional activation factor.

60. The method of claims 59, wherein said transcriptional activation factor is an AraC family polypeptide.

61. The method of claim 59, wherein said transcriptional activation factor is a MarA family polypeptide.

62. The method of any one of claims 1-61, wherein said transcription factor modulating compound is a transcription factor inhibiting compound.

63. The method of any one of claims 1-62, wherein said transcription factor is prokaryotic.

64. The method of claim 61, wherein said MarA family polypeptide is MarA, SoxS, or Rob.

65. The method of claim 1, wherein said microbial cell is selected from the group consisting of *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, *Pseudomonas acidovorans*, *Pseudomonas alcaligenes*, *Pseudomonas putida*, *Stenotrophomonas maltophilia*, *Burkholderia cepacia*, *Aeromonas hydrophilia*, *Escherichia coli*, *Citrobacter freundii*, *Salmonella typhimurium*, *Salmonella typhi*, *Salmonella paratyphi*, *Salmonella enteritidis*, *Shigella dysenteriae*, *Shigella flexneri*, *Shigella sonnei*, *Enterobacter cloacae*, *Enterobacter aerogenes*, *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Serratia marcescens*, *Francisella tularensis*, *Morganella morganii*, *Proteus mirabilis*, *Proteus vulgaris*, *Providencia alcalifaciens*, *Providencia rettgeri*, *Providencia stuartii*, *Acinetobacter calcoaceticus*, *Acinetobacter haemolyticus*, *Yersinia enterocolitica*, *Yersinia pestis*, *Yersinia pseudotuberculosis*, *Yersinia intermedia*, *Bordetella pertussis*, *Bordetella parapertussis*, *Bordetella bronchiseptica*, *Haemophilus influenzae*, *Haemophilus parainfluenzae*, *Haemophilus haemolyticus*, *Haemophilus*

parahaemolyticus, *Haemophilus ducreyi*, *Pasteurella multocida*, *Pasteurella haemolytica*, *Branhamella catarrhalis*, *Helicobacter pylori*, *Campylobacter fetus*, *Campylobacter jejuni*, *Campylobacter coli*, *Borrelia burgdorferi*, *Vibrio cholerae*, *Yibrio parahaemolyticus*, *Legionella pneumophila*, *Listeria monocytogenes*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Gardnerella vaginalis*, *Bacteroides fragilis*, *Bacteroides distasonis*, *Bacteroides* 3452A homology group, *Bacteroides vulgatus*, *Bacteroides ovalus*, *Bacteroides thetaiotaomicron*, *Bacteroides uniformis*, *Bacteroides eggerthii*, *Bacteroides splanchnicus*, *Clostridium difficile*, *Mycobacterium tuberculosis*, *Mycobacterium avium*, *Mycobacterium intracellulare*, *Mycobacterium leprae*, *Corynebacterium diphtheriae*, *Corynebacterium ulcerans*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Enterococcus faecalis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus saprophyticus*, *Staphylococcus intermedius*, *Staphylococcus hyicus subsp. hyicus*, *Staphylococcus haemolyticus*, *Staphylococcus hominis*, and *Staphylococcus saccharolyticus*.

66. A method for identifying a transcription factor modulating compound, comprising:

contacting a microbial cell with a test compound, wherein said microbial cell comprises a selective marker under the direct control of a transcription factor responsive element and a transcription factor, under conditions which allow interaction of the compound with the microbial cell; and

measuring the ability of said test compound to affect said microbial cell, such that transcription factor modulating compounds are identified.

67. The method of claim 66, wherein the responsive element is activated by a transcription factor.

68. The method of claim 67, wherein the responsive element is selected from the group consisting of: *marO*, *fum*, *inaA*, *galT*, and *micF*.

69. The method of claim 66, wherein the responsive element is repressed by a transcription factor.

70. The method of claim 69, wherein the responsive element is selected from the group consisting of *fecA*, *purA*, and *guaB*.

71. The method of claim 66, wherein the selective marker is selected from the group consisting of ccdB, kan, cat, bla, purA, GuaB, and URA3.

72. The method of claim 66, wherein the ability of said test compound to modulate the activity of a transcription factor is demonstrated by the ability of the test compound to promote the *in vitro* or *in vivo* growth or survival of said microbial cell.

73. The method of claim 66, wherein the ability of said test compound to modulate the activity of a transcription factor is demonstrated by the ability of the compound to decrease the *in vitro* or *in vivo* growth or survival said cell in an animal model of infection.

74. The method of claim 66, wherein the method is used to identify a transcription factor agonist.

75. The method of claim 66, wherein the method is used to identify a transcription factor antagonist.

76. The method of claim 66, wherein said transcription factor is expressed from a plasmid.

77. The method of claim 76, wherein said transcription factor is expressed under the control of an inducible promoter.

78. The method of claim 77, wherein said inducible promoter is selected from the group consisting of: trp, tac, tet, and GAL1.

79. The method of claim 76, wherein said transcription factor is expressed under the control of a constitutive promoter.

80. The method of claim 66, wherein said transcription factor is a MarA family polypeptide.

81. The method of claim 66, wherein said microbial cell comprises a chromosomal deletion of at least one gene encoding a transcription factor.

82. The method of claim 66, wherein said transcription factor responsive element comprises a Marbox domain.

83. The method of claim 66, wherein said transcription factor comprises at least one HTH domain.

84. The method of claim 66, wherein said transcription factor is prokaryotic.

85. The method of claim 80, wherein said transcription factor is MarA.

86. A method for identifying a transcription factor modulating compound *in vivo* or *in vitro*, comprising:

contacting a microbial cell comprising: 1) a selective marker under the control of a transcription factor responsive element and 2) a transcription factor, with a test compound under conditions which allow interaction of the compound with the microbial cell; and

measuring the ability of the test compound to affect the growth or survival of the microbial cell as an indication of whether the compound modulates the activity of a transcription factor, wherein the inactivation of the transcription factor leads to a decrease in cell survival.

87. A method for identifying a transcription factor modulating compound *in vivo* or *in vitro*, comprising:

contacting a microbial cell comprising: 1) a selective marker under the control of a transcription factor responsive element and 2) a transcription factor, with a test compound under conditions which allow interaction of the compound with the microbial cell; and

measuring the ability of said test compound to affect the growth or survival of the microbial cell as an indication of whether the compound modulates the activity of a transcription factor, wherein the activation of the transcription factor leads to a decrease in cell survival.

88. A method for identifying a transcription factor modulating compound *in vivo* or *in vitro*, comprising:

contacting a microbial cell comprising: 1) a selective marker under the control of a transcription factor responsive element and 2) a transcription factor, with test compound under conditions which allow interaction of the compound with the microbial cell; and

measuring the ability of said test compound to affect the growth or survival of the microbial cell as an indication of whether said test compound modulates the activity of a transcription factor, wherein the inactivation of the transcription factor leads to an increase in cell survival.

89. A method for identifying a transcription factor modulating compound *in vivo* or *in vitro*, comprising:

contacting a microbial cell comprising: 1) a selective marker under the control of a trascription factor responsive element and 2) a transcription factor, with a compound under conditions which allow interaction of the compound with the microbial cell; and

measuring the ability of the compound to affect the growth or survival of the microbial cell as an indication of whether the compound modulates the activity of a transcription factor, wherein the activation of the transcription factor leads to an increase in cell survival.

90. A method for identifying a transcription factor modulating compound *in vivo* or *in vitro*, comprising:

contacting a microbial cell comprising: 1) a chromosomal deletion in a guaB or purA gene, 2) heterologous guaB or purA gene under the control of its natural promoter, and 3) a transcription factor, with a test compound under conditions which allow interaction of the compound with the microbial cell; and

measuring the ability of the compound to affect the growth or survival of the microbial cell as an indication of whether the compound modulates the activity of a transcription factor, wherein the ability of a compound to modulate the activity of a transcription factor leads to an increase in cell growth.

91. A transcription factor modulating compound identified by the method of any one of claims 66, 86, 87, 88, 89, or 90.

92. The method of any one of claims 86-90, wherein said transcription factor modulating compound affects the growth or development of gram negative bacteria.

93. The method of any one of claims 86-90, wherein said transcription factor modulating compound affects the growth or development of prokaryotic bacteria.

94. The method of any one of claims 86-90, wherein said transcription factor is prokaryotic.

95. The method of any one of claims 86-90, wherein said transcription factor modulating compound affects the growth or development of gram positive bacteria.

96. The method of claim 95, wherein said gram positive bacteria are *Enterococcus*, *Staphylococcus*, *Clostridium* or *Streptococcus*.

97. The method of claim 95, wherein said transcription factor modulating compound affects the growth or development of bacteria from the family *Enterobacteriaceae*.

98. The method of any one of claims 86-90, wherein said transcription factor modulating compound affects the growth or development of *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, *Pseudomonas acidovorans*, *Pseudomonas alcaligenes*, *Pseudomonas putida*, *Stenotrophomonas maltophilia*, *Burkholderia cepacia*, *Aeromonas hydrophilia*, *Escherichia coli*, *Citrobacter freundii*, *Salmonella typhimurium*, *Salmonella typhi*, *Salmonella paratyphi*, *Salmonella enteritidis*, *Shigella dysenteriae*, *Shigella flexneri*, *Shigella sonnei*, *Enterobacter cloacae*, *Enterobacter aerogenes*, *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Serratia marcescens*, *Francisella tularensis*, *Morganella morganii*, *Proteus mirabilis*, *Proteus vulgaris*, *Providencia alcalifaciens*, *Providencia rettgeri*, *Providencia stuartii*, *Acinetobacter calcoaceticus*, *Acinetobacter haemolyticus*, *Yersinia enterocolitica*, *Yersinia pestis*, *Yersinia pseudotuberculosis*, *Yersinia intermedia*, *Bordetella pertussis*, *Bordetella parapertussis*, *Bordetella bronchiseptica*, *Haemophilus influenzae*, *Haemophilus parainfluenzae*, *Haemophilus haemolyticus*, *Haemophilus parahaemolyticus*, *Haemophilus ducreyi*, *Pasteurella multocida*, *Pasteurella haemolytica*, *Branhamella catarrhalis*, *Helicobacter pylori*, *Campylobacter fetus*, *Campylobacter jejuni*, *Campylobacter coli*, *Borrelia burgdorferi*, *Vibrio cholerae*, *Yibrio parahaemolyticus*, *Legionella pneumophila*, *Listeria monocytogenes*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Gardnerella vaginalis*, *Bacteroides fragilis*, *Bacteroides distasonis*, *Bacteroides* 3452A homology group, *Bacteroides vulgatus*, *Bacteroides ovalus*, *Bacteroides thetaiotaomicron*, *Bacteroides uniformis*, *Bacteroides eggerthii*, *Bacteroides splanchnicus*, *Clostridium difficile*,

Mycobacterium tuberculosis, *Mycobacterium avium*, *Mycobacterium intracellulare*, *Mycobacterium leprae*, *Corynebacterium diphtheriae*, *Corynebacterium ulcerans*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Enterococcus faecalis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus saprophyticus*, *Staphylococcus intermedius*, *Staphylococcus hyicus subsp. hyicus*, *Staphylococcus haemolyticus*, *Staphylococcus hominis*, and *Staphylococcus saccharolyticus*.

99. The method of any one of claims 86-90, wherein said transcription factor modulating compound is a nucleic acid molecule.

100. The method of any one of claims 86-90, wherein said transcription factor modulating compound is an antisense or sense oligonucleotide.

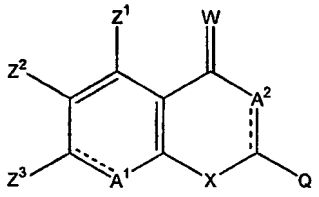
101. The method of any one of claims 86-90, wherein said transcription factor modulating compound is a small molecule.

102. The method of any one of claims 86-101, wherein said transcription factor is a MarA family polypeptide.

103. The method of any one of claims 86-101, wherein said transcription factor is a AraC family polypeptide.

104. A kit for identifying a transcription factor modulating compound which modulates the activity of a transcription factor comprising a microbial cell comprising: 1) a selective marker under the control of a transcription factor responsive element and 2) a transcription factor.

105. A pharmaceutical composition comprising an effective amount of a transcription factor modulating compound, and a pharmaceutically acceptable carrier, wherein said transcription factor modulating compound is of the formula (II):



(II)

wherein

W is O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^4, O, or S;

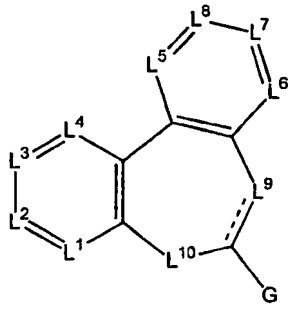
A^2 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, Z^4 and Z^5 are each independently hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, or cyano;

Z^3 is hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, nitro, cyano, carbonyl, or thiocarbonyl;

Q is an aromatic or heterocyclic moiety, and pharmaceutically acceptable salts thereof.

106. A pharmaceutical composition comprising an effective amount of an transcription factor modulating compound, and a pharmaceutically acceptable carrier, wherein said transcription factor modulating compound is of the formula (III):



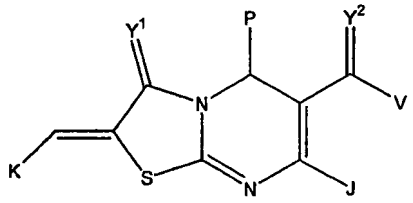
(III)

wherein

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen; and

L^1, L^2, L^3, L^4, L^5, L^6, L^7, L^8, L^9, and L^{10} are each independently oxygen, substituted or unsubstituted nitrogen, sulfur and or substituted or unsubstituted carbon, and pharmaceutically acceptable salts thereof.

107. A pharmaceutical composition comprising an effective amount of a transcription factor modulating compound, and a pharmaceutically acceptable carrier, wherein said transcription factor modulating compound is of the formula (IV):



(IV)

wherein

Y^1 and Y^2 are each oxygen or sulfur;

J is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, cyano, nitro, amino, or halogen;

V is substituted or unsubstituted alkyl, alkenyl, alkynyl, alkoxy, alkylamino, or alkylthio;

P and K are each independently substituted or unsubstituted aryl, and pharmaceutically acceptable salts thereof.

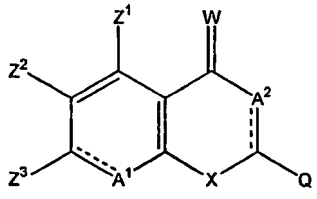
108. The pharmaceutical composition of any one of claims 105-107, further comprising an antibiotic.

109. A pharmaceutical composition of any one of claims 105-107, wherein said effective amount is effective to treat a biofilm associated state in said subject.

110. The pharmaceutical composition of claim 109, wherein said biofilm associated state is selected from the group consisting of middle ear infections, cystic fibrosis, osteomyelitis, acne, dental cavities, endocarditis, and prostatitis.

111. A method of inhibiting a biofilm, comprising administering a composition comprising a transcription factor modulating compound, such that said biofilm is inhibited.

112. The method of claim 111, wherein said transcription factor modulating compound is of the formula (II):



(II)

wherein

W is O or S;

X is O, S, or C, optionally linked to Q;

A^1 is C-Z^4, O, or S;

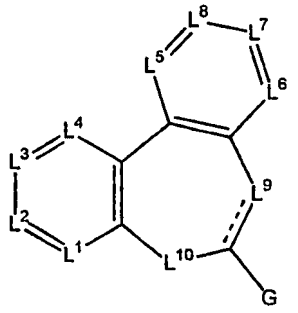
A^2 is C-Z^5, or N-Z^5;

Z^1, Z^2, Z^3, Z^4 and Z^5 are each independently hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, or cyano;

Z^3 is hydrogen, alkoxy, hydroxy, halogen, alkyl, alkenyl, alkynyl, aryl, heterocyclic, amino, nitro, cyano, carbonyl, or thiocarbonyl;

Q is an aromatic or heterocyclic moiety, and pharmaceutically acceptable salts thereof.

113. The method of claim 111, wherein said transcription factor modulating compound is of the formula (III):



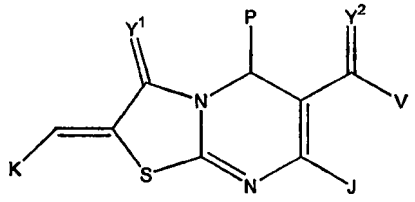
(III)

wherein

G is substituted or unsubstituted aromatic moiety, heterocyclic, alkyl, alkenyl, alkynyl, hydroxy, cyano, nitro, amino, carbonyl, or hydrogen; and

L^1, L^2, L^3, L^4, L^5, L^6, L^7, L^8, L^9, and L^{10} are each independently oxygen, substituted or unsubstituted nitrogen, sulfur and or substituted or unsubstituted carbon, and pharmaceutically acceptable salts thereof.

114. The method of claim 111, wherein said transcription factor modulating compound is of the formula (IV):



(IV)

wherein

Y^1 and Y^2 are each oxygen or sulfur;

J is hydrogen, substituted or unsubstituted alkyl, alkenyl, alkynyl, cyano, nitro, amino, or halogen;

V is substituted or unsubstituted alkyl, alkenyl, alkynyl, alkoxy, alkylamino, or alkylthio;

P and K are each independently substituted or unsubstituted aryl, and pharmaceutically acceptable salts thereof.

115. The method of claim 111, wherein said transcription factor modulating compound is a compound of Table 4 or Table 5.

116. The method of any one of claims 111-115, wherein said composition further comprises a surfactant.

117. The method of claim 116, wherein said surfactant is Sodium Dodecyl Sulfate; Quaternary Ammonium Compounds; alkyl pyridinium iodides; Tween 80, Tween 85, Triton X-100; Brij 56; biological surfactants; Rhamnolipid, Surfactin, Visconsin, or sulfonates.

118. The method of claim 117 wherein said biofilm development is diminished by the administration of said composition.

119. A method of inhibiting the formation of a biofilm, comprising administering a transcription factor modulating compound, such that the formation of said biofilm is inhibited.

120. A method for cleaning and disinfecting contact lenses comprising administering a composition comprising an acceptable carrier and a transcription factor modulating compound, such that said contact lenses are cleaned and disinfected.

121. A method of treating medical indwelling devices comprising administering a composition comprising a transcription factor modulating compound, such that said medical indwelling devices are treated.

122. The method of claim 121, wherein said device is selected from the group consisting of catheters, orthopedic devices and implants.

123. A method for treating or preventing a biofilm associated state in a subject, comprising administering to said subject an effective amount of a transcription factor modulating compound, such that said biofilm associated state in said subject is treated.

124. The method of claim 123, wherein said biofilm associated state is selected from the group consisting of middle ear infections, cystic fibrosis, osteomyelitis, acne, dental cavities, endocarditis, and prostatitis.

125. The method of claim 123, further comprising administering a pharmaceutically acceptable carrier.

126. The method of claim 123, wherein said subject is a mammal.

127. The method of claim 126, wherein said mammal is a human.

128. The method of claim 123, wherein said subject is immunocompromised.

129. A method of identifying a transcription factor modulating compounds, comprising:

obtaining the structure of said transcription factor;

using an appropriate program to identify a scaffold which have an interaction energy score of –20 or less with a portion of said transcription factor, such that transcription factor modulating compounds are identified.

130. The method of claim 129, wherein said transcription factor is a MarA family polypeptide.

131. The method of claim 129, wherein said scaffold has an interaction energy score of –40 or less.

132. The method of claim 131, wherein said scaffold has an interaction energy score of –60 or less.

133. The method of claim 129, wherein said transcription factor is MarA.

134. The method of claim 133, wherein said portion of MarA is selected from the group consisting of about residue 42 to about residue 50, about residue 54 to about residue 62, about residue 55 to about residue 65, about residue 15 to about residue 25, about residue 14 to about residue 25, about residue 24 to about residue 35, about residue 76 to about residue 83, and about residue 106 to about residue 112, of SEQ ID NO. 2.

135. The method of claim 129-132, wherein said transcription factor is Rob.

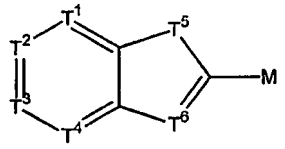
136. The method of claim 135, wherein said portion of Rob is selected from the group consisting of from about residue 37 to about residue 45, about residue 43 to about residue 54, about residue 51 to about residue 60, about residue 10 to about residue 20, about residue 9 to about residue 20, about residue 21 to about residue 29, about

residue 66 to about residue 77, and about residue 101 to about residue 107, of SEQ ID NO. 4.

137. The method of any one of claims129-136, wherein said method further comprises chemically modifying said scaffold.

138. A transcription factor modulating compound identified by any one of the methods of claims 129-137.

139. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and a transcription factor modulating compound, wherein said compound is of the formula (V):



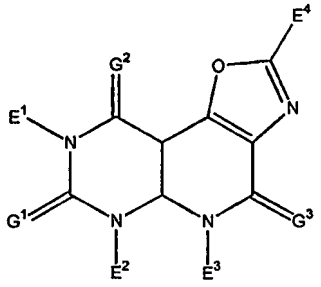
(V)

wherein

T^1, T^2, T^3, T^4, T^5, and T^6 are each independently substituted or unsubstituted carbon, oxygen, substituted or unsubstituted nitrogen, or sulfur;

M is hydrogen, alkyl, alkenyl, heterocyclic, alkynyl, or aryl, or pharmaceutically acceptable salts thereof.

140. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and a transcription factor modulating compound, wherein said compound is of the formula (VI):



(VI)

wherein

G^1, G^2, and G^3 are each independently O, S, substituted or unsubstituted nitrogen, or substituted or unsubstituted carbon;

E^1, E^2, and E^3 are each independently hydrogen, alkyl, alkenyl, alkynyl, aryl, aralkyl, or acyl; and

E^4 is alkyl, alkenyl, alkynyl, aryl, halogen, cyano, amino, nitro, or acyl, and pharmaceutically acceptable salts thereof.

141. A method for preventing a bacterial associated state in a subject, comprising administering to said subject an effective amount of a transcription factor modulating compound, such that the bacterial associated state in said subject is prevented.

142. The method of claim 141, wherein said subject is a human.

143. The method of claim 141, wherein said transcription factor modulating compound is a MarA family polypeptide inhibitor.

144. The method of claim 141, wherein said transcription factor modulating compound is a AraC family polypeptide inhibitor.

Multiple sequence alignment of AraC-XylS family members.

```
AARP_PROST/22-120       SEILVWIEGNLTNR.....LSLDDIAQHSGYTKWHLQRVFRKIVGMPLGEYIRRRRI
ADA_ECOLI/85-183        DKITHACRLLEQETp.....VTLEALADQVAMSPFHLHRLFKATTGMTPKAWQQAWRA
ADA_MYCTU/87-185        ARAMRLIADGTVDR.....DGVSGLAAQLGYTIRQLERLLQAVVGAGPLALARAQRM
ADA_SALTY/94-183        ----------LEQEt.....pVTLAFLAQAVAMSPFHLHRLFKASTGMTPKGWQQAWRA
ADAA_BACSU/102-200      DLITEYIDKNFTEK.....LTLESLADICHGSPYHMHRTFKKIKGITLVEYIQQVRV
ADIY_ECOLI/149-246      DSVYQIIESDIHKD.....WNLSMVASCLCLSPSLLKKKLKSENT-SYSQIITTCRM
AGGR_ECOLI/164-261      DKVRNTIEKDLSKR.....WTLAIIADEFNVSEITIRKRLESEYI-TFNQILMQSRM
APPY_ECOLI/139-236      CKITGIISENIERQ.....WHLKDIAELIYTSESLIKKRLRDEGT-SFTEILRDTRM
ARAC_CITFR/180-279      RDACQYISDHLADSn.....FDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI
ARAC_ECOLI/180-279      REACQYISDHLADSn.....FDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI
ARAC_ERWCH/186-284      IEACQFITSNLAGE.....LRIDEVARHVCLSPSRLAHLFREQVGINILRWREDQRV
ARAC_SALTY/180-279      RDACQYISDHLADSh.....FDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI
ARAL_STRAT/202-300      ASALTFLHRDPAHS.....WTVAELASAAAVSRSTLAARFKATVGQGPLEYLTRWRI
ARAL_STRLI/202-300      ATALTCLHRDPARS.....WTVADLADTAAVSRSTLAARFKATVGQGPLEYLTRWRI
CAFR_YERPE/8-107        NSIIQYIEENLESKf.....INIDCLVLYSGFSRRYLQISFKEYVGMPIGTYIRVRRA
CELD_ECOLI/168-274      DDVPQWLKSTVEKMhdkeqfsESALENMVALSAKSQEYLTRATQRYYGKTPMQIINEIRI
CFAD_ECOLI/164-261      DKVRNVIEKDLSRK.....WTLGIIADAFNVSEITIRKRLESENT-NFNQILMQLRM
CSVR_ECOLI/166-263      DKVRGVIEKDLSRK.....WTLAIIADVFNVSEITIRKRLESEDT-NFNQILMQSRM
ENV_ECOLI/149-246       DSVCRIIQSDIQHY.....WNLRIVASSLCLSPSLLKKKLKNENT-SYSQIVTECRM
EUTR_ECOLI/243-344      SRAREYVLENMSEP.....VTVLDLCNQLHVSRRTLQNAFHAILGIGPNAWLKRIRL
EUTR_SALTY/243-344      SRAREYVLENMSEP.....LTVLDLCNQLHVSRRTLQNAFHAILGIGPNAWLKRIRL
EXSA_PSEAE/171-269      ERLQLFMEKHYLNE.....WKLSDFSREFGMGLTTFKELFGSVYGVSPRAWISERRI
FAPR_ECOLI/154-251      ERIVTLLFSDLTRK.....WKLSDIAEEMHISEISVRKRLEQECL-NFNQLILDVRM
FEAR_ECOLI/199-299      QKVVTLIDDNIREEi.....LRPEWIAGETGMSVRSLYRMFADKGL-VVAQYIRNRRL
GADX_ECO27/145-242      TRVCTVINNNIAHE.....WTLARIASELLMSPSLLKKKLREEGT-SYSQLLTECRM
GADX_ECO57/145-242      TRVCTVINNNIAHE.....WTLARIASELLMSPSLLKKKLREEET-SYSQLLTECRM
GADX_ECOLI/145-242      TRVCTVINNNIAHE.....WTLARIASELLMSPSLLKKKLREEET-SYSQLLTECRM
GLXA_RHIME/223-321      LAVLEKMETAIERP.....LDRTAMARLAGVSPRHLDRLFREHRGTGFLDTYREIRL
HRPB_RALSO/375-477      RRAYRYIIENIERSd.....LTTKEVAAHINVTERALQLAFKSAVGMSPSSVIRRMRL
INVF_SALTY/212-210      YWLVGYLLAQSTSG.....NTMRMLGEDYGVSYTHFRRLCSRALGGKAKSELRNWRM
LACR_STAXY/274-272      QHAVDFINTNYQKH.....ITVEDVAKSVNITRSHLYKLFKKNLGCSPKEYLTYIRM
LCRF_YERPE/167-265      ERLQKFMEENYLQG.....WKLSKFAREFGMGLTTFKELFGTVYGISPRAWISERRI
LUMQ_PHOLE 148-246      VLIDNYIEQHLQKK.....ISVAELSSVAFLAQSQFYALFKSQMGITPHQYVLRKRL
MARA_ECOLI/14-112       HSILDWIEDNLESP.....LSLEKVSERSGYSKWHLQRMFKKETGHSLGQYIRSRKM
MARA_SALTY/14-112       HSILDWIEDNLESP.....LSLEKVSERSGYSKWHLQRMFKKETGHSLGQYIRSRKM
MELR_ECOLI/194-292      SQMLGFIAENYDQA.....LTINDVAEHVKLNANYAMGIFQRVMQLTMKQYITAMRI
MMSR_PSEAE/201-299      DGLHAYMREHLHAR.....LELERLAAFCNLSKFHFVSRYKAITGRTPIQHFLHLKI
MSMR_STRMU/176-274      NQVKKIIHSQYGSS.....LRVNDIAKKLNLSRSYLYKIFRKSTNLSIKEYILQVRM
MXIE_SHIFL/99-199       YHLVLYLLRTIEKEk.....eVRIKSLTEHYGVSEAYFRSLCRKALGAKVKEQLNTWRL
MXIE_SHISO/99-199       YHLVLYLLRTIEKEk.....eVRIKSLTEHYGVSEAYFRSLCRKALGAKVKEQLNTWRL
ORUR_PSEAE/241-338      TRVRRLLLARPGDF.....PDLEQAAREIHTSGRSLRRHLSSLGT-TYQQVLDDVRK
PCHR_PSEAE/201-296      HAARDLLVGALQEP.....PSLDTLASRVGMNPRKLTAGFRKVFGASVFGYLQEYRL
PERA_ECO27/168-265      DRVIKVIELDISKN.....WKLGDVSSSMFMSDSCLRKQLNKENL-TFKKIMLDIKM
POCR_SALTY/195-293      KKALRYIDAHLSDD.....LRLEDVASHVYLSPYYFSKLFKKYQGIGFNAWVNRQRM
PQRA_PROVU/7-107        NDILKWLETQLQRNe.....gIKIDTIANKSGYSKWHLQRIFKDFKGCTLGEYVRKRRL
RAFR_PEDPE/176-274      NLAVSYLQENYSTG.....CTIMDLCHYINLSRSYLYTLFKTHANTSPQKLLTKLRL
RAMA_ENTCL/9-107        DTIVEWIDDNLHQP.....LRIEDIARHAGYSKWHLQRLFLQYKGESLGRYIRERKL
RAMA_KLEPN/9-107        DTIVEWIDDNLHQP.....LRIDDIARHAGYSKWHLQRLFLQYKGESLGRYIRERKL
RHAR_ECOLI/209-307      DKLITRLAASLKSP.....FALDKFCDEASCSERVLRQQFRQQTGMTINQYLRQVRV
RHAR_SALTY/179-277      DKLITALANSLECP.....FALDAFCQQEQCSERVLRQQFRAQTGMTINQYLRQVRI
RHAS_ECOLI/174-272      NLLLAWLEDHFADE.....VNWDAVADQFSLSLRTLHRQLKQQTGLTPQRYLNRLRL
RHAS_SALTY/174-272      NQLMAWLEDHFAEE.....VCWEAVAEQFSLSLRTLHRQLKQHTGLTPQRYLNRLRL
RHRA_RHIME/210-310      ASIKMRVEQNLANGs.....FSITDVAEAERITPRAIQKFFSREGT-TFSRYVLGRRL
RNS_ECOLI/164-261       DKVRNLIEKDLSRK.....WTLGIIADAFNASEITIRKRLESENT-NFNQILMQLRM
```

Figure 1A

```
ROB_ECOLI/8-106          RDLLIWLEGHLDQP.....LSLDNVAAKAGYSKWHLQRMFKDVTGHAIGAYIRARRL
SOXS_ECOLI/7-105         QDLIAWIDEHIDQP.....LNIDVVAKKSGYSKWYLQRMFRTVTHQTLGDYIRQRRL
SOXS_SALTY/7-105         QTLIEWIDEHIDQP.....LNIDVVAKKSGYSKWYLQRMFRTVTHQTLGEYIRQRRL
TCPN_VIBCH/172-269       EKISCLVKSDITRN.....WRWADICGELRTNRMILKKELESRGV-KFRELINSIRI
TETD_ECOLI/31-129        KDVLLWIEHNLDQS.....LLLDDVANKAGYTKWYFQRLFKKVTGVTLASYIRARRL
THCR_RHOER/227-328       RLAVDYLEAHAQQP.....LTVAQVARNVGVSVRSLQVGFQNSLGTTPMRQLKIIRM
URER_ECOLI/171-268       QAITHLITQEPQKK.....WHLDDVAKALFTTPSTLRRHLNREGV-SFRQLLLDVRM
URER_PROMI/171-268       QAITHLITQDPQRK.....WHLEDVAKTLYTTPSTLRRHLSKEGV-SFCQLLLDVRI
VIRF_SHIDY/161-258       DQIRKIVEKNIEKR.....WRLSDISNNLNLSEIAVRKRLESEKL-TFQQILLDIRM
VIRF_YEREN/167-265       ERLQKFMEENYLQG.....WKLSKFAREFGMGLTTFKELFGTVYGISPRAWISERRI
VIRS_MYCTU/236-334       ERVVGLARRLLPTGq.....CSAEAIADQLDMHPRTLQRRLAAEGL-RCHDLIERERR
XYLR_ECOLI/288-386       IQAMHYIRNHACKG.....IKVDQVLDAVGISRSNLEKRFKEEVGETIHAMIHAEKL
XYLR_HAEIN/288-386       IQAMHYIRHRACHR.....IKVGQVLDHLETCRSNLEQRFKNEMNKTIHQVIHEEKI
XYLS_PSEPU/214-315       ERVVQFIEENLKRN.....ISLERLAELAMMSPRSLYNLFEKHAGTTPKNYIRNRKL
XYS1_PSEPU/214-315       ERVVQFIEENLKRN.....ISLERLAELAMMSPRSLYNLFEKHAGTTPKNYIRNRKL
XYS2_PSEPU/39-140        ERVVQFIEENVKRS.....ISLEQLAELAIMSPRSLYTMFEKHTGTTPMNYIRNRKL
XYS3_PSEPU/214-315       ERVVQFIEDNLKQS.....ISLERLAELAIMSPRSLYTLFEKHAGTTPKNYIRNRKL
XYS4_PSEPU/214-315       ERVVQFIEENLKRN.....ISLERLAELAIMSPRSLYTLFEKHAGTTPKNYIRNRKL
Y4FK_RHISN/318-417       LKAEAFMRENLTNP.....VTIEDLAAAARCTPRALQRMFRTYRGGSPMSVLCNYRL
YA52_HAEIN/194-295       KRLNTALLAILQQPqn....dWHIEQLAELATMSRANFIRIFQQHIGMSPGRFLTKVRL
YBBB_BACSU/166-264       EKTKHYIETHADTK.....ITLAQLSQMAGISAKHYSESFKKWTGQSVTEFITKTRI
YBCM_ECOLI/165-267       SRCYNLLLSEPGTK.....WTANKVARYLYISVSTLHRRLASEGV-SFQSILDDVRL
YCGK_ALTCA/67-163        QNAMLYIENNYFND.....INIDTVAFSVGVSRSYLVKQFKLATNKTINNRIIEVRI
YD95_MYCTU/242-343       RGITALVRSKLFRDsg....lfPTFTDVAGELDMHPRTLRRRLAEEGT-SFRALLGEARS
YDEO_ECOLI/137-233       GKVRNIVNMKPAHP.....WKLKDICDCLYISESLLKKKLKQEQT-TFSQILLDARM
YDIP_ECOLI/183-281       KDILFYLNNNYREK.....ITLEQLSKKFRASVSYICHEFTKEYRISPINYVIQRRM
YEAM_ECOLI/158-258       PKIRTMVEMMAKGPve....wGALGQWAGFFAMSERNLARLIVKETGLSFRQWRQQLQL
YFIF_BACSU/192-289       TEVKLHIKDNLSQP.....LKLTDVASHFHISGRHLSRLFAAELGVSYSEFVQNEKI
YHIW_ECOLI/139-236       GKVERLISFDIAKR.....WYLRDIAERMYTSESLIKKKLQDENT-CFSKILLASRM
YIDL_ECOLI/197-295       EKLIATLHASLQQR.....WSVADMAATIPCSEAWLRRLFLRYTGKTPKEYYLDARL
YIJO_ECOLI/172-270       EAIRDYIDERYASA.....LTRESVAQAFYISPNYLSHLFQKTGAIGFNEYLNHTRL
YISR_BACSU/183-281       WEAARYLQEHYKEK.....TTIKDLSLALHYHQDYVSRCMQQVLGVTPAQYTNRVRM
YKGA_ECOLI/19-117        QQLLEWIECNLEHP.....ISIEDIAQKSGYSRRNIQLLFRNFMHVPLGEYIRKRRL
YKGD_ECOLI/177-278       PRLGAVIQQMLEMPgh....aWTVESLASIAHMSRASFAQLFRDVSGTTPLAVLTKLRL
YMCR_STRLA/184-281       DPLLRAVVVSLEAG.....RSVTATADSVGLGARQLHRRSLAAFGYGPKTLARVLRM
YPDC_ECOLI/184-282       HSICNWVQDNYAQP.....LTRESVAQFFNITPNHLSKLFAQHGTMRFIEYVRWVRM
YQHC_ECOLI/213-311       SRVLKRIENKYTEN.....LSVEQLAAEANMSVSAFHHNFKSVTSTSPLQYLKNYRL

AARP_PROST/22-120        CEAAKELQTTNL...QVIDIALKYQFDSQQSFAKRFKAYLGISPSLYRLS
ADA_ECOLI/85-183         RRLRESLAKGE-...SVTTSILNAGFPDSSSYYRKADETLGMTAKQFRHG
ADA_MYCTU/87-185         QTARVLIETTNL...PFGDVAFAAGFSSIRQFNDTVRLACDGTPTALRAR
ADA_SALTY/94-183         RRLREALAKGE-...PITAAIYRAGFPDSSSYYRHADQTLGMTAKQFRKG
ADAA_BACSU/102-200       HAAKKYLIQTNK...AIGDIAICVGIANAPYFITLFKKKTGQTPAEFRQM
ADIY_ECOLI/149-246       RYAVNELMMDGK...NISQVSQSCGYNSTSYFISVFKDFYGMTPLHYVSQ
AGGR_ECOLI/164-261       SKAALLLLDNSY...QISQISNMIGFSSTSYFIRLFVKHFGITPKQFLTY
APPY_ECOLI/139-236       RYAKKLITSNSY...SINVVAQKCGYNSTGYFICAFKDYYGVTPSHYFEK
ARAC_CITFR/180-279       SQAKLLLSTTRM...PIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG
ARAC_ECOLI/180-279       SQAKLLLSTTRM...PIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG
ARAC_ERWCH/186-284       IRAKLLLQTTQE...SIANIGRVVGYDDQLYFSRVFKKRVGVSPSDFRRR
ARAC_SALTY/180-279       SQAKLLLSTTRM...PIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG
ARAL_STRAT/202-300       ELTARQLREGSA...PLAAIAHSVGYGSESALSVAFKRVLGMNPGDYRKH
ARAL_STRLI/202-300       ELAARQLREGNA...TLASIAHSVGYGSESALSVAFKRVLGMPPGDYRKH
CAFR_YERPE/8-107         SRAAALLRLTRL...TIIEISAKLFYDSQQTFTREFKKIFGYTPRQYRMI
CELD_ECOLI/168-274       NFAKKQLEMTNY...SVTDIAFEAGYSSPSLFIKTFKKLTSFTPKSYRKK
CFAD_ECOLI/164-261       SKAALLLLENSY...QISQISNMIGISSASYFIRVFNKHYGVTPKQFFTY
CSVR_ECOLI/166-263       SKAALLLLENSY...QISQISNMIGISSASYFIRIFNKHFGVTRSSFLII
```

Figure 1B

```
ENVY_ECOLI/149-246      RYAVQMLIMDNK...NITQVAQLCGYSSTSYFISVFKAFYGLTPLNYLAK
EUTR_ECOLI/243-344      NAVRRELISPWSqsaTVKDAAMQWGFWHLGQFATDYQQLFSEKPSLTLHQ
EUTR_SALTY/243-344      NAVRRELISPWSqsaTVKDAAMQWGFWHLGQFATDYQQLFAEKPSLTLHQ
EXSA_PSEAE/171-269      LYAHQLLLNSDM...SIVDIAMEAGFSSQSYFTQSYRRRFGCTPSRSRQG
FAPR_ECOLI/154-251      NQAAKFIIRSDH...QIGMIASLVGYTSVSYFIKTFKEYYGVTPKKFEIG
FEAR_ECOLI/199-299      DFCADAIRHAADd.eKLAGIGFHWGFSDQSHFSTVFKQRFGMTPGEYRRK
GADX_ECO27/145-242      QRALQLIVIYGV...SIKRVAVSCGYHSVSYFIYVFRNYYGMTPTEYQER
GADX_ECO57/145-242      QRALQLIVIHGF...SIKRVAVSCGYHSVSYFIYVFRNYYGMTPTEYQER
GADX_ECOLI/145-242      QRALQLIVIHGF...SIKRVAVSCGYHSVSYFIYVFRNYYGMTPTEYQER
GLXA_RHIME/223-321      RHARRLLQQSPL...SIPEIAYATGFSSPAHFSNAFKRLFSQTPGSLRRR
HRPB_RALSO/375-477      EGIRSDLLDSERnpsNIIDTASRWGIRSRSALVKGYRKQFNEAPSETIWR
INVF_SALTY/112-210      AQSLLNSVEGHE...NITQLAVNHGYSSPSHFSSEIKELIGVSPRKLSNI
LACR_STAXY/174-272      YHASQLLIHTST...LISDISRQVGYKDPLLFSKNFTKHFEISASEYRHH
LCRF_YERPE/167-265      LYAHQLLLNGKM...SIVDIAMEAGFSSQSYFTQSYRRRFGCTPSQARLT
LUMQ_PHOLE/148-246      DLAKQLIAERQK...PLSQVAQLCGFSSQSSFSQAFRRLYGMSPTRYQFF
MARA_ECOLI/14-112       TEIAQKLKESNE...PILYLAERYGFESQQTLTRTFKNYFDVPPHKYRMT
MARA_SALTY/14-112       TEIAQKLKESNE...PILYLAERYGFESQQTLTRTFKNYFDVPPHKYRIT
MELR_ECOLI/194-292      NHVRALLSDTDK...SILDIALTAGFRSSSRFYSTFGKYVGMSPQQYRKL
MMSR_PSEAE/201-299      EYACQLLDSSDQ...SVARVGQAVGYDDSYYFSRLFSKVMGLSPSAYRQR
MSMR_STRCO/176-274      KRSQYLLENPKL...SIAEISNSVGFSDSLAFSKAFKNYFGKSPSKFRKE
MXIE_SHIFL/99-199       VNGLLDVFLHNQ...TITSAAMNNGYRSTSHFSNEIKTRLGFSARELSNI
MXIE_SHISO/99-199       VNGLLDVFLHNQ...TITSAAMNNGYASTSHFSNEIKTRLGFSARELSNI
ORUF_PSEAE/241-338      RLALQYLTTTQL...PLYEIALLLGFNDSSNFRRAFRKWTGKLPSDYREA
PCHR_PSEAE/201-296      REAHRMLCDEEA...NVSTVAYRVGYS-PAHFSIAFRKRYGISPSEIR--
PERA_ECO27/168-265      KHASLFLRTTDK...NIDEISCLVGFNSTSYFIKVFKEYYNTTPKKYNGV
POCR_SALTY/195-293      VSARELLCHSDW...SIASIARNLGFSQTSYFCKVFRQTYQVTPQAYRQQ
PQRA_PROVU/7-107        LEAAKSLQEKDM...SILDIALMYGFSSQATFTRIFKKHFNTTPAKFREN
RAFR_PEDPE/176-274      EDAKQRLSTSNN...SVQSIANMVGYKDSFTFSKAFKRYSGASPSYYRKS
RAMA_ENTCL/9-107        LLAARDLRESDE...RVYEICLRYGFESQQTFTRIFTRTFHQPPGAYRKE
RAMA_KLEPN/9-107        LLAARDLRDTDQ...RVYDICLKYGFDSQQTFTRVFTRTFNQPPGAYRKE
RHAR_ECOLI/209-307      CHAQYLLQHSRL...LISDISTECGFEDSNYFSVVFTRETGMTPSQWRHL
RHAR_SALTY/179-277      CHAQYLLQHSPL...MISEISMQCGFEDSNYFSVVFTRETGMTPSQWRHL
RHAS_ECOLI/174-272      MKARHLLRHSEA...SVTDIAYRCGFSDSNHFSTLFRREFNWSPRDIRQG
RHAS_SALTY/174-272      IKARHLLRHSDH...SVTEIAYRCGFGDSNHFSTLFRREFNWSPRDIRQG
RHRA_RHIME/210-310      SLAKSLILAEGEa.tSISQIAYNVGFNDLSYFNRTFRSRYGVRPSDLRRL
RNS_ECOLI/164-262       SKAALLLLENSY...QISQISNMIGISSASYFIRIFNKHYGVTPKQFFTY
ROB_ECOLI/8-106         SKSAVALRLTAR...PILDIALQYRFDSQQTFTRAFKKQFAQTPALYRRS
SOXS_ECOLI/7-105        LLAAVELRTTER...PIFDIAMDLGYVSQQTFSRVFRRQFDRTPSDYRHR
SOXS_SALTY/7-105        LLAAVELRTTER...PIFDIAMDLGYVSQQTFSRVFRREFDRTPSDYRHR
TCPN_VIBCH/172-269      SYSISLMKTGEF...KIKQIAYQSGFASVSYFSTVFKSTMNVAPSEYLFM
TETD_ECOLI/31-129       TKAAVELRLTKK...TILEIALKYQFDSQQSFTRRFKYIFKVTPSYYRRN
THCR_RHOER/227-328      QKARKDLLRADPaseGVTEIAQRWGFLHVGRFAGEYKQTFGVSPSEDLRT
URER_ECOLI/171-268      GMALNYLTFSNY...SVFQISHRCGFGSNAYFCDVFKRKYNMTPSQFRLQ
URER_PROMI/171-268      PIALNYLTFSNY...SVFQISHRCGFGSNAYFCDAFKRKYGMTPSQFRTQ
VIRF_SHIDY/161-258      HHAAKLLLNSQS...YINDVSRLIGISSPSYFIRKFNEYYGITPKKFYLY
VIRF_YEREN/167-265      LYAHQLLLNGKM...SIVDIAMEAGFSSQSYFTQSYRRRFGCTPSQARLT
VIRS_MYCTU/236-334      AQAARYLAQPGL...YLSQIAVLLGYSEQSAINRSCRRWFGMTPRQYRAY
XYLR_ECOLI/288-386      EKARSLLISTTL...SINEISQMCGYPSLQYFYSVFKKAYDTTPKEYRDV
XYLR_HAEIN/288-386      SRAKNLLQQTDI...SIKEITEICGYPSIQYFYSVFKKEFEMTPKEFRLN
XYLS_PSEPU/214-315      ESIRACLNDPSAnvrSITEIALDYGFLHLGRFAENYRSAFGELPSDTLRQ
XYS1_PSEPU/214-315      ESIRACLNDPSAnvrSITEIALDYGFLHLGRFAENYRSAFGELPSDTLRQ
XYS2_PSEPU/39-140       ECVRACLSNPTTnirSITEVALDYGFLHLGRFAEKYRSTFGELPSDTLSL
XYS3_PSEPU/214-315      ECIRARLSDPNAnvrSVTEMALDYGFFHTGRFAENYRSTFGELPSDTLRR
XYS4_PSEPU/214-315      ECIRARLSDPNAnvrSVTEMALDYGFFHTGRFAENYRSTFGELPSDTLRR
Y4FK_RHISN/318-417      AAAHGAIKAGRAg..SITELALNLQFSNPGRFSVLYKSAYGLSPSSALRF
YA52_HAEIN/194-295      QSAAFLLKQSQQ...SVLAIALEVGYQSEAHFCKVFKNYYQLSPSQYRKS
YBBB_BACSU/166-264      TKAKRIMAKSNC...KLKEIAHQTGYQDEFYFSRIFKKYTGCSPTSYMKK
YBCM_ECOLI/165-262      NNALSAIQTTVK...PISEIARENGYKCPSRFTERFHNRFNITPRSIRKA
```

Figure 1C

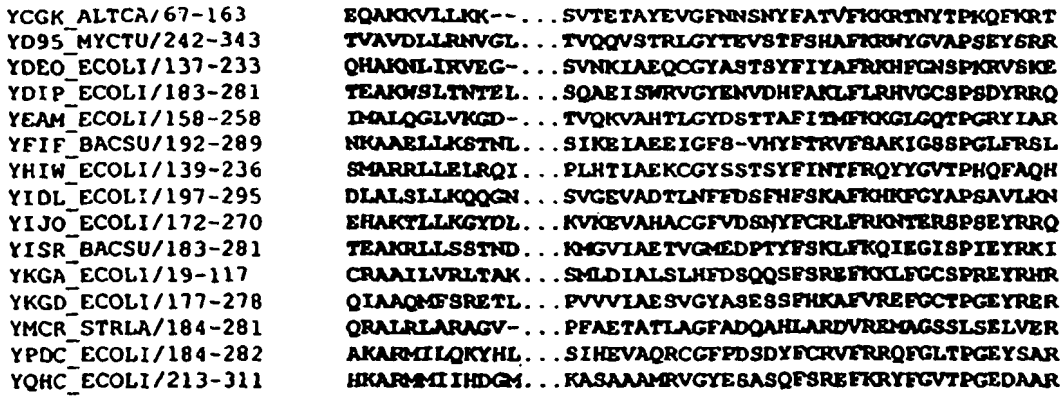

YCGK_ALTCA/67-163 EQAKKVLLKK--...SVTETAYEVGFNNSNYFATVFKKRTNYTPKQFKRT
YD95_MYCTU/242-343 TVAVDLLRNVGL...TVQQVSTRLGYTEVSTFSHAFKRWYGVAPSEYGRR
YDEO_ECOLI/137-233 QHAIQNLIRVEG-...SVNKIAEQCGYASTSYFIYAFRKHFGNSPKRVSKE
YDIP_ECOLI/183-281 TEAKWSLTNTEL...SQAEISWRVGYENVDHFAKLFLRHVGCSPSDYRRQ
YEAM_ECOLI/158-258 DMALQGLVKGD-...TVQKVAHTLGYDSTTAFIRMFKKGLGQTPGRYIAR
YFIF_BACSU/192-289 NKAAELLKSTNL...SIKEIAEEIGFS-VHYFTRVFSAKIGSSPGLFRSL
YHIW_ECOLI/139-236 SMARRLLELRQI...PLHTIAEKCGYSSTSYFINTFRQYYGVTPHQFAQH
YIDL_ECOLI/197-295 DLALSLLKQQGN...SVGEVADTLNFFDSFHFSKAFKHKFGYAPSAVLKN
YIJO_ECOLI/172-270 EHAKTLLKGYDL...KVKEVAHACGFVDSNYFCRLFRKNTERSPSEYRRQ
YISR_BACSU/183-281 TEAKRLLSSTND...KMGVIAETVGMEDPTYFSKLFKQIEGISPIEYRKI
YKGA_ECOLI/19-117 CRAAILVRLTAK...SMLDIALSLHFDSQQSFSREFKKLFGCSPREYRHR
YKGD_ECOLI/177-278 QIAAQMFSRETL...PVVVIAESVGYASESSFHKAFVREFGCTPGEYRER
YMCR_STRLA/184-281 QRALRLARAGV-...PFAETATLAGFADQAHLARDVREMAGSSLSELVER
YPDC_ECOLI/184-282 AKARMILQKYHL...SIHEVAQRCGFPDSDYFCRVFRRQFGLTPGEYSAR
YQHC_ECOLI/213-311 HKARMMIIHDGM...KASAAAMRVGYESASQFSREFKRYFGVTPGEDAAR

Figure 1D

Multiple sequence alignment of PROSITE PS00041, HTH_AraC family 1.

```
AARP_PROST/72-114       RIceAAkeLqtt...nlqVidIAlkyQFdsqqsFakrFKaylGiSP
ADA_MYCTU/137-179       RMqtARvlIett...nlpFgdVAfaaGFssirqFndtVRlacDgTP
ADAA_BACSU/152-194      RVhaAKkyLiqt...nkaIgdIAicvGIanapyFitlFKkktGqTP
ADIY_ECOLI/198-240      RMryAVneLmmd...gknIsqVSqscGYnstsyFisvFKdfyGmTP
AGGR_ECOLI/213-255      RMskAAllLldn...syqIsqISnmiGFsstsyFirlFVkhfGiTP
APPY_ECOLI/188-230      RMryAKklItsn...sysInvVAqkcGYnstsyFicaFKdyyGvTP
ARAC_CITFR/231-273      RIsqAKllLstt...rmpIatVGrnvGFddqlyFsrvFKkctGaSP
ARAC_ECOLI/231-273      RIsqAKllLstt...rmpIatVGrnvGFddqlyFsrvFKkctGaSP
ARAC_ERWCH/236-278      RVirAKllLqtt...qesIanIGrvvGYddqlyFsrvFRkrvGvSP
ARAC_SALTY/231-273      RIsqAKllLstt...rmpIatVGrnvGFddqlyFsrvFKkctGaSP
ARAL_STRAT/252-294      RIelTArqLreg...sapLaaIAhsvGYgsesaLsvaFKrvlGmNP
ARAL_STRLI/252-294      RIelAArqLreg...natLasIAhsvGYgsesaLsvaFKrvlGmPP
CELD_ECOLI/226-268      RInfAKkqLemt...nysVtdIAfeaGYssqslFiktFKkltSfTP
CFAD_ECOLI/211-255      QLrmSKeaLlll...e.nsyqIsqISnmiGIssasyFirvFNkhyGvTP
CSVR_ECOLI/215-257      RMskAAllLlen...syqIsqISnmiGIssasyFiriFNkhfGvTR
ENVY_ECOLI/198-240      RMryAVqmLlmd...nknItqVAqlcGYsstsyFisvFKafyGlTP
EUTR_ECOLI/293-338      RLnaVRreLispwsqsmtVkdAAmqwGFwhlgqFatdYQqlfSeKP
EUTR_SALTY/293-338      RLnaVRreLispwsqsatVkdAAmqwGFwhlgqFatdYQqlfAeKP
EXSA_PSEAE/221-263      RIlyAHqlLlns...dmsIvdIAmeaGFssqsyFtqsYRrrfGcTP
FAPR_ECOLI/203-245      RMnqAAkfIirs...dhqIgmIAslvGYtsvsyFiktFKeyyGvTP
GADX_ECO27/194-236      RMqrALqlIviy...gvsIkrVAvscGYhsvsyFiyvFRnyyGmTP
GADX_ECO57/194-236      RMqrALqlIvih...gfsIkrVAvscGYhsvsyFiyvFRnyyGmTP
GADX_ECOLI/194-236      RMqrALqlIvih...gfsIkrVAvscGYhsvsyFiyvFRnyyGmTP
GLXA_RHIME/273-315      RLrhARrlLqqs...plsIpeIAyatGFsspahFsnaFKrlfSqTP
HRPB_RALSO/426-471      RLegIRsdLldsernpsnIidTAsrwGIrsrsaLvkgYRkqfNeAP
INVF_SALTY/162-204      RMaqSLlnSveg...henItqLAvnhGYsspshFsseIKeliGvSP
LACR_STAXY/224-266      RMyhASqlLiht...stlIsdISrqvGYkdpllFsknFTkhfEiSA
LCRF_YERPE/217-259      RIlyAHqlLlng...kmsIvdIAmeaGFssqsyFtqsYRrrfGcTP
LUMQ_PHOLE/198-240      RLdlAKqlIaer...qkpLsqVAqlcGFssqssFsqaFRrlyGmSP
MARA_ECOLI/64-106       KMteIAqkLkes...nepIlyLAeryGFssqqtLtrtFKnyfDvPP
MARA_SALTY/64-106       KMteIAqkLkes...nepIlyLAeryGFssqqtLtrtFKnyfDvPP
MELR_ECOLI/244-286      RInhVRalLsdt...dksIldIAltaGFrssrFystFGkyvGmSP
MMSR_PSEAE/251-293      KIeyACqlLdss...dqsVarVGqavGYddsyyFsrlFSkvmGlSP
MSMR_STRMU/226-268      RMkrSQylLenp...klsIaeISnsvGFsdslaFskaFKnyfGkSP
MXIE_SHIFL/151-193      RLvnGLldVflh...nqtItsAAmnnGYrstshFsneIKtrlGfSA
MXIE_SHISO/151-193      RLvnGLldVflh...nqtItsAAmnnGYastshFsneIKtrlGfSA
PCHR_PSEAE/251-292      RLreAHrmLcds...eanVstVAyrvGYsp.ahFsiaFRkryGiSP
PERA_ECO27/217-259      KMkhASlfLrtt...dknIdeISclvGFnstsyFikvFKeyyNtTP
POCR_SALTY/245-287      RMvsARelLchs...dwsIasIArnlGFsqtsyFckvFRqtyQvTP
PQRA_PROVU/59-101       RLleAAksLqek...dmsIldIAlmyGFssqatFtriFKkhfNtTP
RAFR_PEDPE/226-268      RLedAKqrLsts...nnsVqsIAnmvGYkdsftFskaFKrysGaSP
RAMA_KLEPN/59-101       KLllAArdLrdt...dqrVydIClkyGFdsqqtFtrvFTrtfNqPP
RHAR_ECOLI/259-301      RVchAQylLqhs...rllIsdIStacGFedsnyFsvvFTretGmTP
RHAR_SALTY/229-271      RIchAQylLqhs...plmIseISmqcGFedsnyFsvvFTretGmTP
RHAS_ECOLI/224-266      RImkARhlLrhs...easVtdIAyrcGFsdsnhFstlFRrefNwSP
RHAS_SALTY/224-266      RLikARhlLrhs...dhsVteIAyrcGFgdsnhFstlFRrefNwSP
RNS_ECOLI/211-255       QLrmSKeaLlll...e.nsyqIsqISnmiGIssasyFiriFNkhyGvTP
SOXS_ECOLI/57-99        RLllAAveLrtt...erpIfdIAmdlGYvsqqtFsrvFRrqfDrTP
SOXS_SALTY/57-99        RLllAAveLrtt...erpIfdIAmdlGYvsqqtFsrvFRrefDrTP
TCPN_VIBCH/221-263      RIsySIslMktg...efkIkqIAyqsGFasvsyFstvFKstmNvAP
THCR_RHOER/277-322      RMqkARkdLlradpasegVteIAqrwGFlhvgrFageYKqtfGvSP
URER_ECOLI/220-262      RMgmALnyLtfs...nysVfqIShrcGFgsnayFcdvFKrkyNmTP
```

Figure 2A

```
URER_PROMI/220-262      RIpiALnyLtfs...nysVfqISbrcGPgsnayPcdaPKrkyGmTP
VIRF_SHIDY/210-252      RMhhAAklLlns...qsyIndVSrliGIespsyPirkPMeyyGiTP
VIRF_YEREN/217-259      RIlyAHqlLlng...kmsIvdIAmeaGPssqsyPtqsYRrrfGcTP
XYLR_ECOLI/338-380      KLekARslList...tlsInsISqmcGYpslqyPysvPKkayDtTP
XYLR_HAEIN/338-380      KIsrAKnlLqqt...disIkeITeicGYpsiqyPysvPKkefEmTP
XYLS_PSEPU/264-309      KLesIRacLndpsanvrsIteIAldyGPlhlgrPaenYRsafGeLP
XYS1_PSEPU/264-309      KLesIRacLndpsanvrsIteIAldyGPlhlgrPaenYRsafGeLP
XYS2_PSEPU/89-134       KLecVRacLsnpttnirsIteVAldyGPlhlgrPaekYRstfGeLP
XYS3_PSEPU/264-309      KLecIRarLsdpnanvrsVteMAldyGPfhtgrPaenYRstfGeLP
XYS4_PSEPU/264-309      KLecIRarLsdpnanvrsVteMAldyGPfhtgrPaenYRstfGeLP
Y4FK_RHISN/368-411      RLaaAHgaIkagr..agsIteLAlnlQPsnpgrPsvlYKsayGlSP
YA52_HAEIN/247-289      RLqsAAflLkqs...qqsVlsIAlevGYqseahPckvPKnyyQlSP
YBBB_BACSU/216-258      RItkAKrlMaks...nckLkeIAhqtGYqdefyPsriPKkytGcSP
YCGK_ALTCA/117-157      RIeqAKkvLlk.....ksVteTAyevGPnnsnyPatvPKkrtNyTP
YDEO_ECOLI/186-227      RMqhAKnlIrve....gsVnkIAeqcGYastsyPiyaPRkhfGnSP
YEAM_ECOLI/211-252      QLimALqgLvkg....dtVqkVAhtlGYdsttaPitmPKkglGqTP
YFIF_BACSU/242-283      KInkAAelLkst...nlsIkeIAeeiGPsv.hyPtrvPSakiGsSP
YHIW_ECOLI/188-230      RMsmARrlLelr...qipLhtIAekcGYsstsyPintPRqyyGvTP
YIDL_ECOLI/247-289      RLdlALslLkqq...gnsVgeVAdtlNPfdsfhPskaPKhkfGyAP
YIJO_ECOLI/222-264      RLehAKtlLkgy...dlkVkeVAhacGPvdsnyPcrlPRkntErSP
YISR_BACSU/233-275      RMteAKrlLsst...ndkMgvIAetvGMedptyPsklPKqieGiSP
YMCR_STRLA/234-275      RMqrALrlArag....vpPaeTAtlaGPadqahLardVRemaGsSL
YPDC_ECOLI/234-276      RMakARmiLqky...hlsIheVAqrcGPpdsdyPcrvPRrqfGlTP
ADA_ECOLI/86-128        KIthACrlLeqet..pvtLeaLAdqvAMsp.fhLhrlPKattGmTP
ADA_ECOLI/136-177       RArrLResLakg....esVttSIlnaGPpdsssYyrkADetlGmTA
ADA_SALTY/86-128        KIacACrlLeqet..pvtLafLAqavAMsp.fhLhrlPKastGmTP
ADA_SALTY/136-177       RArrLReaLakg....epItaAIyraGPpdsssYyrhADqtlGmTA
CAPR_YERPE/56-101       RVrrASraAallrltrltIieISaklPYdsqqtPtrePKkifGyTP
ROB_ECOLI/55-100        RArrLSksAvalrltarpIldIAlqyRPdsqqtPtraPKkqfAqTP
TETD_ECOLI/78-123       RArrLTkaAvelrltkktIleIAlkyQPdsqqsPtrrPKyifKvTP
```

Figure 2B

```
Multiple sequence alignment PS01124, HTH_ARAC_FAMILY_2.

AARP_PROST/22-120       SEILVWIEGNLTNR.....LSLDDIAQHSGYTKWHLQRVFRKIVGMPLGEYIRRRI
ADA_ECOLI/85-183        DKITHACRLLEQETp.....VTLEALADQVAMSPFHLHRLFKATTGMTPKAWQDARA
ADA_MYCTU/87-185        ARAMRLIADGTVDR.....DGVSGLAAQLGYTIRQLERLLQAVVGAGPLALARAQRM
ADA_SALTY/94-183        ----------LEQEt.....pVTLAFLAQAVAMSPFHLHRLFKASTGMTPKQWQQAMRA
ADAA_BACSU/102-200      DLITEYIDKNFTEK.....LTLESLADICHGSPYHMHRTFKKIKGITLVEYIQQVRV
ADIY_ECOLI/149-246      DSVYQIIESDIHKD.....WNLSMVASCLCLSPSLLKKKLKSENT-SYSQIITTCRM
AGGR_ECOLI/164-261      DKVRNTIEKDLSKR.....WTLAIIADEFNVSEITIRKRLESEYI-TFNQILMQSRM
APPY_ECOLI/139-216      CKITGIISFNIERQ.....WHLKDIAELIYTSESLIKKALRDEGT-SFTEILRDTRM
ARAC_CITFR/180-279      RDACQYISDHLADSn.....FDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI
ARAC_ECOLI/180-279      REACQYISDHLADSn.....FDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI
ARAC_ERWCH/186-284      IEACQFITSNLAGE.....LRIDEVARHVCLGPSRLAHLFREQVGINILRWREDQRV
ARAC_SALTY/180-279      RDACQYISDHLADSh.....FDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI
ARAL_STRAT/202-300      ASALTFLHRDPAHS.....WTVAELASAAAVSRSTLARRFKATVGQSPLEYLTRWRI
ARAL_STRLI/202-300      ATALTCLHRDPARS.....WTVADLADTAAVSRSTLARRFKATVGQGPLEYLTRWRI
CAFR_YERPE/8-107        NSIIQYIEENLESKf.....INIDCLVLYSGFSRRYLQISFKEVVGMPIGTYIRVRRA
CELD_ECOLI/168-274      DDVPQWLKSTVEKMhdkeqfseSALENMVALSAKSQEYLTRATQRYYGKTPMQIINEIRI
CFAD_ECOLI/164-261      DKVRNVIEKDLSRK.....WTLGIIADAFNVSEITIRKRLESENT-NFNQILMQLRM
CSVR_ECOLI/166-263      DKVRGVIEKDLSRK.....WTLAIIADVFNVSEITIRKRLESEDT-NFNQILMQSRM
ENVY_ECOLI/149-246      DSVCRIIQSDIQHY.....WNLRIVASSLCLSPSLLKKKLKNENT-SYSQIVTECRM
EUTR_ECOLI/243-344      SRAREYVLENMSEP.....LTVLDLCNQLHVSRRTLQNAFHAILGIGPNAWLKRIRL
EUTR_SALTY/243-344      SRAREYVLENMSEP.....LTVLDLCNQLHVSRRTLQNAFHAILGIGPNAWLKRIRL
EXSA_PSEAE/171-269      ERLQLFMEKHYLNE.....WKLSDFSREFGMGLTTFKELFGSVYGVSPRAWISERRI
FAPR_ECOLI/154-251      ERIVTLLFSDLTRK.....WKLSDIAEEMHISEISVRKRLEQECL-NFNQLILDVRM
FEAR_ECOLI/199-299      QKVVTLIDDNIREEi.....LRPEWIAGETGMSVRSLYRMFADKGL-VVAQYIRNRRL
GADX_ECO27/145-242      TRVCTVINNNIAHE.....WTLARIASELLMSPSLLKKKLREEGT-SYSQLLTECRM
GADX_ECOSJ/145-242      TRVCTVINNNIAHE.....WTLARIASELLMSPSLLKKKLREEET-SYSQLLTECRM
GADX_ECOLI/145-242      TRVCTVINNNIAHE.....WTLARIASELLMSPSLLKKKLREET-SYSQLLTECRM
GLXA_RHIME/223-321      LAVLEKMETAIERP.....LDRTAMARLAGVSPRHLDRLFREHRGTGFLDTIREIRL
HRPB_RALSO/375-477      RRAYRVIIENIERSd.....LTTREVAAHINVTERALQLAFKSAVGMSPSSVIRRMRL
INVF_SALTY/112-210      YWLVGYLLAQSTSG.....NTMRMLGEDYGVSYTHFRRLCSRALGGKAKSELRNWRM
LACR_STAXY/174-272      QHAVDFINTNYQKH.....ITVEDVAKSVNITRSHLYKLFKKNLGCSPKEYLTYIRM
LCRF_YERPE/167-265      ERLQKFMEENYLQG.....WKLSKFAREFGMGLTTFKELFGTVYGISPRAWISERRI
LUMQ_PHOLE/148-246      VLIDNYIEQHLQKK.....ISVAELSSVAFLAQSQFYALFKSQMGITPHQYVLRKRL
MARA_ECOLI/14-112       HSILDWIEDNLESP.....LSLEKVSERSGYSKWHLQRMFKKETGHSLGQYIRSRKM
MARA_SALTY/14-112       HSILDWIEDNLESP.....LSLEKVSERSGYSKWHLQRMFKKETGHSLGQYIRSRKM
```

Figure 3A

```
MELR_ECOLI/194-292      SQMLGFIAENYDQA.....LTINDVAEHVKLNANYAMGIFQRVMQLTMKQVITAMRI
MMSR_PSEAE/201-299      DGLHAVDEHLHAR.....LELERLAAFCNLSKPHFVSRYKAITGRTPIQHFLHLKI
MSMR_STRMU/176-274      NQVKKIIHSQYGSS.....LRVNDIAKKLNLSRSYLYKIFRKSTNLSIKEYILQVRM
MXIE_SHIFL/99-199       YHLVLYLLRTIEKEk.....eVRIKSLTEHYGVSEAYFRSLCRKALGAKVKEQLNTWRL
MXIE_SHISO/99-199       YHLVLYLLRTIEKEk.....eVRIKSLTEHYGVSEAYFRSLCRKALGAKVKEQLNTWRL
ORUR_PSEAE/241-338      TRVRRLLLARPGDF.....PDLEQAARELHTSGRSLRRHLSSLGT-TYQQVLDDVRK
PCHR_PSEAE/201-296      HAARDLLVGALQEP.....PSLDTLASRVGMNPRKLTAGFRKVFGASVPGYLQEYRL
PSRA_ECO27/168-265      DRVIKVIELDISKN.....WKLGDVSSSMFMSDSCLRKQLNKENL-TFKKIMLDIKM
PSRA_ECOLI/168-265      DRVIKVIELDISKN.....WKLGDVSSSMFMSDSCLRKQLNKENL-TFKKIMLDIKM
PSCR_SALTY/195-293      KKALRYIDAHLSDD.....LRLEDVASHVYLSPYYFSKLFKKYQGIGFNAYVNRQRM
PQRA_PROVU/7-107        NDILKWLETQLQRNe.....gIKIDTIANKSGYSKWHLQRIFKDFKGCTLGEYVRGRL
RAFR_PEDPE/176-274      NLAVSYLQENYSTG.....CTIMDLCHYLNLSRSYLYTLFKTHANTSPQKLLTGLRL
RAMA_ENTCL/9-107        DTIVEWIDDNLHQP.....LRIEDIARHAGYSKWHLQRLFLQYKGESLGRYIRERKL
RAMA_KLEPN/9-107        DTIVEWIDDNLHQP.....LRIDDIARHAGYSKWHLQRLFLQYKGESLGRYIRERKL
RHAR_ECOLI/209-307      DKLITRLAASLKSP.....FALDKFCDEASCSERVLRQQFRQQTGMTINQYLRQVRV
RHAR_SALTY/179-277      DKLITRLAASLECP.....FALDAFCQQEQCSERVLRQQFRAQTGMTINQYLRQVRI
RHAS_ECOLI/174-272      NLLLAWLEDHFADE.....VNWDAVADQFSLSLRTLHRQLKQQTGLTPQRYLNRLRL
RHAS_SALTY/174-272      NQLMAWLEDHFAEE.....VCWEAVAEQFSLSLRTLHRQLKQHTGLTPQRYLNRLRL
RHPA_RHIME/210-310      ASIKMRVEQNLANGe.....FSITDVAEAERITPRAIQKFFSREGT-TFSRYVLGRRL
RNS_ECOLI/164-261       DKVRNLIEKDLSRK.....WTLGIIADAFNASEITIRKRLESENT-NFNQILMQLRM
ROB_ECOLI/8-106         RDLLIWLEGHLDQP.....LSLDNVAAKAGYSKWHLQRMFKDVTGHAIGAYIRARRL
SOXS_ECOLI/7-105        QDLIAWIDEHIDQP.....LN1DVVAKKSGYSKWYLQRMFRTVTHQTLGDYIRQRRL
SOXS_SALTY/7-105        QTLIEWIDEHIDQP.....LNIDVVAKKSGYSKWYLQRMFRTVTHQTLGEYIRQRRL
TCPN_VIBCH/172-269      EKISCLVKSDITRN.....WRWADICGELRTNRMILKKELESSRGV-KFRELINSIRI
TETD_ECOLI/31-129       KDVLLWIEHNLDQS.....LLLDDVANKAGYTKWYFQRLFKKVTGVTLASYIRARRL
THCR_RHOER/227-328      RLAVDYLEAHAQQP.....LTVAQVARNVGVSVRSLQVGFQNSLGTTPMRQLKIRM
URER_ECOLI/171-268      QAITHLITQEPQKK.....WHLDDVAKALFTTPSTLRRHLNREGV-SFRQLLLDVRM
URER_PROMI/171-268      QAITHLITQDPQRK.....WHLEDVAKTLYTTPSTLRRHLSKEGV-SFCQLLLDVRI
VIRF_SHIDY/161-258      DQIRKIVEKNIEKR.....WRLSDISNNLNLSEIAVRKRLRSEKL-TFQQILLDIRM
VIRF_YEREN/167-265      ERLQKFNEENYLQG.....WKLSKPAREFGMGLTTFKELFGTVYGISPRAWISERRI
VIRS_MYCTU/236-334      ERVVGLARRLLPTGq.....CSAEAIADQLDMHPRTLQRRLAAEGL-RCHDLIERERR
XYLR_ECOLI/288-386      IQAMRYIRNHACKG.....IKVDQVLDAVGISRSNLEKRFKEEVGETIHAMIHAEKL
XYLR_HAEIN/288-386      IQAMRYIRHRACHR.....IKVGQVLDHLETSRSNLEQRFKNENMKTIHQVIHEEKI
XYLS_PSEPU/214-315      ERVVQFIEENLKRN.....ISLERLAELAMSPRSLYNLFEKHAGTTPKQYIRNRKL
XYS1_PSEPU/214-315      ERVVQFIEENLKRN.....ISLERLAELAMSPRSLYNLFEKHAGTTPKQYIRNRKL
XYS2_PSEPU/39-140       ERVVQFIEENVKRS.....ISLEQLAELAMSPRSLYTMFEKHTGTTPMQYIRNRKL
XYS3_PSEPU/214-315      ERVVQFIEDNLKQS.....ISLERLAELAMSPRSLYTLFEKHAGTTPKQYIRNRKL
XYS4_PSEPU/214-315      ERVVQFIEENLKRN.....ISLERLAELAMSPRSLYTLFEKHAGTTPKQYIRNRKL
Y4FK_RHISN/318-417      LKAEAFNRENLTNP.....VTIEDLAAAARCTPRALQRMFRTYRGGSPMSVLQNYRL
YAS2_HAEIN/194-295      KRLNTALIAILQQPqn.....dWHIEQLAELATMSRANFIRIFQQHIGMSPGRFLTKVRL
```

Figure 3B

```
YBBB_BACSU/166-264   EKTKHYIETHADTK.....ITLAQLSQWAGISAKHYSESFKKWTGQSVTEFITKTRI
YBCM_ECOLI/165-263   CRCYNLIISEPGTK.....WTANKVARVLYISVSTLHPRLASEGV-SFQSILDDVRL
YCGK_ALTCA/67-163    QNAMLYIENNYFND.....INIDTVAFSVGVSRSYLVKQFKLATNKTINNRIIEVRI
YD95_MYCTU/242-343   RGITALVRSKLFRDsg....lfPTFTDVAGELDMHPRTLRRRLAEEGT-SFRALLGEARS
YDEO_ECOLI/137-233   GKVRNIVNMKPAHP.....WKLKDICDCLYISESLLKKKLKQEQT-TFSQILLDARM
YDIP_ECOLI/183-281   KDILFYLNNNYREK.....ITLEQLSKKFRASVSYICHEFTKEYRISPINYVIQRRM
YEAM_ECOLI/158-258   PKIRTMVEMMAKGPve....wGALGQNAGFFAMSERNLARLIVKETGLSFRQWRQQLQL
YFIF_BACSU/192-289   TEVKLHIKDNLSQP.....LKLTDVASHFHISGRHLSRLFAAELGVSYSEFVQNEKI
YHIW_ECOLI/139-236   GKVERLISFDIAKR.....WYLRDIAERMTTSESLIKKKLQDENT-CFSKILLASRM
YIDL_ECOLI/197-295   EKLIATLHASLQQR.....WSVADMAATIPCSEAWLRRLPLRYTGKTPKEYYLDARL
YIJO_ECOLI/172-270   EAIRDYIDERYASA.....LTRESVAQAFYISPNYLSHLFQKTGAIGFNEYIANTRL
YISR_BACSU/183-281   WEAARYLQEHYKEK.....TTIKDLSLALHYHQDYSRCMQQVLGVTPAQYTNRVRM
YKGA_ECOLI/19-117    QQLLEWIECNLEHP.....ISIEDIAQKSGYSRNIQLLFRNFMHVPLGEYIRKRRL
YKGD_ECOLI/177-278   PRLGAVIQQMLEMPgh....aWTVESLASIAHMSRASFAQLFRDVSGTTPLAVLTKLRL
YMCR_STRLA/184-281   DPLLRAVVVSLEAG.....RSVTATADSVGLGARQLHRRSLAAFGYGPKTLARVLRM
YPDC_ECOLI/184-282   HSICNWVQDNYAQP.....LTRESVAQFFNITPNMLSKLFAQHGTMRFIEYVRWVRM
YQHC_ECOLI/213-311   SRVLKRIENKYTEN.....LSVEQLAAEANMSVSAFHHNFKSVTSTSPLQYLKQYRL

AARP_PROST/22-120    CEAAKELQTTNL...QVIDIALKYQFDSQQSFAKRFKAYLGISPSLYRLS
ADA_ECOLI/85-183     RRLRESLAKGE-...SVTTSILNAGFPDSSSYYRKADETLGMTAKQFRHG
ADA_MYCTU/87-185     QTARVLIETTNL...PFGDVAFAAGFSSIRQFNDTVRLACDGTPTALRAR
ADA_SALTY/94-183     RRLREALAKGE-...PITAAIYRAGFPDSSSYYRHADQTLGMTAKQFRKG
ADAA_BACSU/102-200   HAAKKYLIQTNK...AIGDIAICVGIANAPYFITLFKKKTGQTPARFRQM
ADIY_ECOLI/149-246   RYAVNELMMDGK...NISQVSQSCGYNSTSYFISVFKDFYGMTPLHYVSQ
AGGR_ECOLI/164-261   SKALLLLDNSY...QISQISNMIGFSSTSYFIRLFVKHFGITIPKQFLTY
APPY_ECOLI/139-236   RYAKKLITSNSY...SINVVAQKCGYNSTSYFICAFKDYYGVTPSHYFEK
ARAC_CITFR/180-279   SQAKLLLSTTRM...PIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG
ARAC_ECOLI/180-279   SQAKLLLSTTRM...PIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG
ARAC_ERWCH/186-284   IRAKLLLQTTQE...SIANIGRVVGYDDQLYFSRVFRKRVGVSPSDFRRR
ARAC_SALTY/180-279   SQAKLLLSTTRM...PIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG
ARAL_STRAT/202-300   ELTARQLREGSA...PLAAIAHSVGYGSESALSVAFKRVLGMNPGDYRKH
ARAL_STRLI/202-300   ELAARQLREGNA...TLASIAHSVGYGSESALSVAFKRVLGMPPGDYRKH
CAFR_YERPE/8-107     SRAAALLRLTRL...TIIEISAKLFYDSQQTFTREFKKIFGYTPRQYRMI
CELD_ECOLI/168-274   NFAKKQLEMTNY...SVTDIAFEAGYSSPSLFIKTFKKLTSFTPKSYRKK
CFAD_ECOLI/164-261   SKAALLLLENSY...QISQISNMIGISSASYFIRVFNKHYGVTPKQFFTY
CSVR_ECOLI/166-263   SKAALLLLENSY...QISQISNMIGISSASYFIRIPNKHFGVTRSSFLII
ENVY_ECOLI/149-246   RYAVQMLMDNK...NITQVAQLCGYSSTSYFISVFKAFYGLTPLNYLAK
```

Figure 3C

```
EUTR_ECOLI/243-344   NAVRRELISPWSqsmTVKDAAMQWGFWHLGQFAIDYQQLFSERPSLTLHQ
EUTR_SALTY/243-344   NAVRRELISPWSqsaTVKDAAMQWGFWHLGQFAIDYQQLFSERPSLTLHQ
EXSA_PSEAE/171-269   LYAHQLLNSDM...SIVDIAMEAGFSSQSYFTQSYRRRFGCTPSRSRQG
FAPR_ECOLI/154-251   NQAAKFIIRSDH...QIGMIASLVGYTSVSYFIKTFKEYYGVTPKKFEIG
FEAR_ECOLI/199-299   DFCADAIRHAADd.eKLAGIGFHWGFSDQSHFSTVFKQRFGMTPGEYRRK
GADX_ECO27/145-242   QRALQLIVIYGV...SIKRVAVSCGYHSVSYFIYVFRNTYGMTPTEYQEK
GADX_ECOS7/145-242   QRALQLIVIHGF...SIKRVAVSCGYHSVSYFIYVFRNTYGMTPTEYQER
GADX_ECOLI/145-242   QRALQLIVIHGF...SIKRVAVSCGYHSVSYFIYVFRNTYGMTPTEYQER
GLXA_RHIME/223-321   RHARRLLQQSPL...SIPEIAYATGFSSPAHFSNAFKRLFSQYPGSLRRR
HRPB_RALSO/375-477   EGIRSDLLDSERnpsNIIDTASRWGIRSRSALVKGYRKQFNEAPSETIWR
INVF_SALTY/112-210   AQSLLNSVEGHE...NITQLAVNHGYSSPSHFSSEIKELLGVSPRKLSNI
LACR_STAXY/174-272   YHASQLLIHTST...LISDISRQVGYKDPLLFSKNFTKHFEISASEYRHH
LCRF_YERPE/167-265   LYAHQLLLNGKM...SIVDIAMEAGFSSQSYFTQSYRRRFGCTPSQARLT
LUMQ_PHOLE/148-246   DLAKQLIAERQK...PLSQVAQLCGFSSQSSFSQAFRRLYGMSPTRYQFP
MARA_ECOLI/14-112    TEIAQKLKESNE...PILYLAERYGFESQQTLTRTFKNYFDVPPHKYRMT
MARA_SALTY/14-112    TEIAQKLKESNE...PILYLAERYGFESQQTLTRTFKNYFDVPPHKYRIT
MELR_ECOLI/194-292   NHVRALLSDTDK...SILDIALTAGFRSSSRFYSTFGKYVGMSPQQYRKL
MMSR_PSEAE/201-299   EYACQLLDSSDQ...SVARVGQAVGYDDSYYFSRLFSKVMGLSPSAYRQR
MSMR_STRMU/176-274   KRSQYLLENPKL...SIAEISNSVGFSDSLAFSKAFKNYFGKSPSKFRKE
MXIE_SHIFL/99-199    VNGLLDVFLHNQ...TITSAAMNNGYASISHFSNEIKTRLGFSARELSNI
MXIE_SHISO/99-199    VNGLLDVFLHNQ...TITSAAMNNGYASISHFSNEIKTRLGFSARELSNI
ORUR_PSEAE/241-338   RLALQYLTTTQL...PLYEIALLLGFNDSSNFRRAFRKWTGKLPSDYREA
PCHR_PSEAE/201-296   REAHRMLCDEEA...NVSTVAYKVGYS-PAHFSIAFKKRYGISPSEIR--
PERA_ECO27/168-265   KHASLFLRTTDK...NIDEISCLVGFNSTSYFIKVFKEYYNTTPKKYNGV
POCR_SALTY/195-293   VSARELLCHSDW...SIASIARNLGFSQTSYFCKVFRQTYQVTPQAYRQQ
PQRA_PROVU/7-107     LEAAKSLQEKDM...SILDIALMYGFSSQATFTKIFKKHFNTTPAKFREN
RAFR_PEDPE/176-274   EDAKQRLSTSNN...SVQSIANMVGYKDSFTFSKAFKRYSGASPSYYRKS
RAMA_ENTCL/9-107     LLAARDLRESDE...RVYEICLRYGFESQQTFTRIFTKQFPQAYRKE
RAMA_KLEPN/9-107     LLAARDLRDTDQ...RVYDICLKYGFDSQQTFTRVFTRTFNQFPGAYRKE
RHAR_ECOLI/209-307   CHAQYLLQHSPL...LISDISTECGFEDSNYFSVVFTRETGMTPSQWRHL
RHAR_SALTY/179-277   CHAQYLLQHSPL...MISEISMQCGFEDSNYFSVVFTRETGMTPSQWRHL
RHAS_ECOLI/174-272   MKARHLLRHSEA...SVTDIAYRCGFSDSNHFSTLFRREFNWSPRDIRQG
RHAS_SALTY/174-272   IKARHLLRHSDH...SVTEIAYRCGFDSNHFSTLFRREFNWSPRDIRQG
RHRA_RHIME/210-310   SLAKSLILAEGEa.tSISQIAYNVGFNDLSYFNRTFRSRYGVTPRSDLRRL
RNS_ECOLI/164-261    SKAALLLLENSY...QISQISNMIGISSASYFIRIFNKHYGVTPKQFFTY
ROB_ECOLI/8-106      SKSAVALRLTAR...PILDIALQYRFDSQQIFTRAFKKQFAQTPALYRRS
SOXS_ECOLI/7-105     LLAAVELRTTER...PIFDIAMDLGYVSQQTFSRVFTRREFDRTPSDYRHR
SOXS_SALTY/7-105     LLAAVELRTTER...PIFDIAMDLGYVSQQTFSRVFTRREFDRTPSDYRHR
TCPN_VIBCH/172-269   SYSISLMKTGEF...KIKQIAYQSGFASVSYFSTVFKSTMNVAPSEYLFM
```

Figure 3D

```
TETD_ECOLI/31-129          TKAAVELRLTKK...TLLEIALKYQPDSQQSFTRRFKYIKVTTSYTRRN
THCR_RHOER/227-328         QKARKDLLRADPaseGVTEIAQRWGPLHVGRFAGEYKQTFQVSDSEDLRT
URER_ECOLI/171-268         GMALNYLTFSNY...SVFQISHRCGPGSNAYFCDVPKRKYRMTPSQFRLQ
URER_PROMI/171-268         PIALNYLTFSNY...SVFQISHRCGPGSNAYFCLAPKRNYGMTPSQFRTQ
VIRF_SHIDY/161-258         HHAAKLLLNSQS...YINDVSKLIGISSPSYTIRKFNEYGITPKKPYLY
VIRP_YEREN/167-265         LYAHQLLLNGKM...SIVDIAMEAGPSSQSIYTQSYRKFGCTPSQARLT
VIRS_MYCTU/236-334         AQAARYLAQPGL...YLSQIAVLLGYSEQSALNRSCRRWFGMTPRQYRAY
XYLR_ECOLI/288-386         EKARSLLISTTL...SINEISQMCGYPSLQYFKSVFKKAYDTTPKEYRDV
XYLR_HAEIN/288-386         SRAKNLLQQTDI...SIKEITEICGYPSIQYFYSVFKKEFEMTPKEFRLN
XYLS_PSEPU/214-315         ESIRACLANDPSAnvrSITEIALDYGFLHLGRFAENYRSAFGELPSDTLRQ
XYS1_PSEPU/214-315         ESIRACLANDPSAnvrSITEIALDYGFLHLGRFAENYRSAFGELPSDTLRQ
XYS2_PSEPU/19-140          ECVRACLSNPTThirSITEVALDYGFLHLGRFAEKYRSTFGELPSDTLSL
XYS3_PSEPU/214-315         ECIRARLSDPNAnvrSVTEMALDYGFFHTGRFAENYRSTFGELPSDTLRR
XYS4_PSEPU/214-315         ECIRARLSDPNAnvrSVTEMALDYGFFHTGRFAENYRSTFGELPSDTLRR
Y4FK_RHISN/318-417         AAAHGAIKAGRAg..SITELALNLQFSNPGRFSVLYKSAYGLSPSSALRF
YA52_HAEIN/194-295         QSAAFLLKQSQQ...SVLAIALEVGYQSEAHFCKVFKQYYQLSPSQYRKS
YBBB_BACSU/166-264         TKAKRLMAKSNC...KLKEIAKQTGYQDEFIFSRIFKKYTGCSPTSYMKK
YBCM_ECOLI/165-262         NNALSAIQTTVK...PISEIARENGYKCPSRFTERFRNRFNITPREIRKA
YCGK_ALTCA/67-163          EQAKKVLLKK-...SVTETAYEVGFRNSNYFATVFKKRTNYTPKQFKRT
YD95_MYCTU/242-343         TVAVDLLRNVGL...TVQQVSTRLGYTEVSTFSHAFKRWTGVAPSEYSRR
YDEO_ECOLI/137-233         QHAKNLIRVEG-...SVNKIAEQCGYASTSYFIYAFRKQFGNSPKRVSKE
YDIP_ECOLI/183-281         TEAKWSLTNTEL...SQAEISWRVGYENVDHFAKLFLHHVGCSPSDYTRRQ
YEAM_ECOLI/158-258         IMALQGLVKGD-...TVQKVAHTLGYDSITAFITMFKKGLGQTPGRYIAR
YFIF_BACSU/192-289         NKAAELLKSTNL...SIKEIAEEIGFS-VHYFTRVFSAKIGSSPGLFRSL
YHIW_ECOLI/139-236         SMARRLLELRQI...PLHTIAEKCGYSSTSYFINTFRQYYGVTPHQFAQH
YIDL_ECOLI/197-295         DLALSLLKQQGN...SVGEVADTLNFFDSFHFSKAFKHKNYTAPSAVLKN
YIJO_ECOLI/172-270         EHAKTLLKGYDL...KVKEVAHACGFVDSNYFCRLFRKNTERSPSEYTRRQ
YISR_BACSU/183-281         TEAKRLLSSTND...KMGVIAETVGMEDPTYFSKLFKQIEMISPIEYRKI
YKGA_ECOLI/19-117          CRAAILVRLTAK...SMLDIALSLHFDSQSFSREFKKLFGCSPREYRHR
YKGD_ECOLI/177-278         QIAAQMFSRETL...PVVVIAESVGYASESSFHKAFTREFGCTPGEYRER
YMCR_STRLA/184-281         QRALRLARAGV-...PFAETATLAGFADQAHLARDVREWAGSSLSELVER
YPDC_ECOLI/184-282         AKARMILQKYHL...SIHEVAQRCGFPDSDYFCRVFARQPGLTPGEYSAS
YQHC_ECOLI/213-311         HKARMMIIHDGM...KASAAAMRVGYESASQFSRLFKRYFGVTPGEDAAR
```

Figure 3E

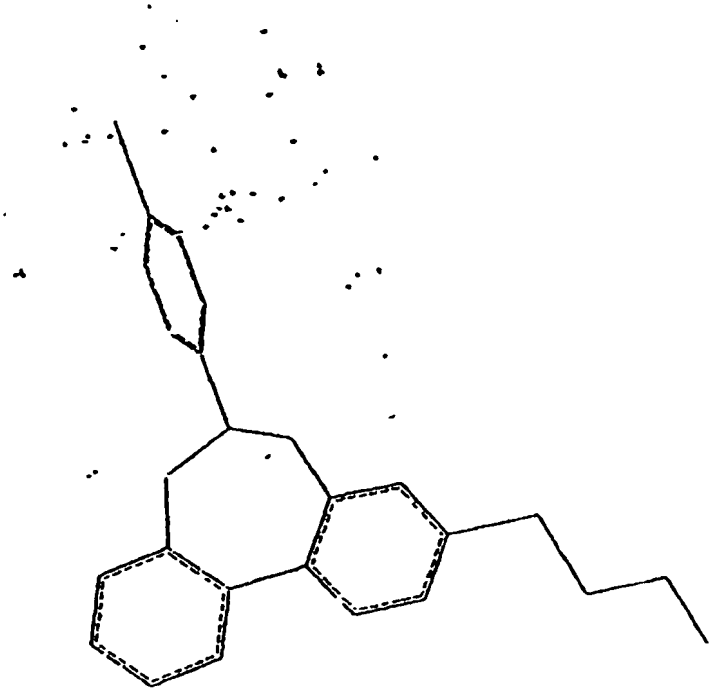


Figure 4

SEQUENCE LISTING

<110> PARATEK PHARMACEUTICALS, INC. et al.

<120> TRANSCRIPTION FACTOR MODULATING COMPOUNDS AND METHODS OF USE THEREOF

<130> PAZ-054CPPC

<150> US 60/288,660
<151> 2001-05-04

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 7878
<212> DNA
<213> Echerichia coli

<220>
<221> CDS
<222> (4124)...(4843)

<400> 1
gttaactgtg gtggttgtca ccgcccatta cacggcatac agctatatcg agccttttgt 60
acaaaacatt gcgggattca gcgccaactt tgccacggca ttactgttat tactcggtgg 120
tgcgggcatt attggcagcg tgattttcgg taaactgggt aatcagtatg cgtctgcgtt 180
ggtgagtacg gcgattgcgc tgttgctggt gtgcctggca ttgctgttac ctgcggcgaa 240
cagtgaaata cacctcgggg tgctgagtat tttctggggg atcgcgatga tgatcatcgg 300
gcttggtatg caggttaaag tgctggcgct ggcaccagat gctaccgacg tcgcgatggc 360
gctattctcc ggcatattta atattggaat cggggcgggt gcgttggtag gtaatcaggt 420
gagtttgcac tggtcaatgt cgatgattgg ttatgtgggc gcggtgcctg cttttgccgc 480
gttaatttgg tcaatcatta tatttcgccg ctggccagtg acactcgaag aacagacgca 540
atagttgaaa ggcccattcg ggccttttt aatggtacgt tttaatgatt tccaggatgc 600
cgttaataat aaactgcaca cccatacata ccagcaggaa tcccatcaga cgggagatcg 660
cttcaatgcc acccttgccc accagccgca taattgcgcc ggagctgcgt aggcttcccc 720
acaaaataac cgccaccagg aaaaagatca gcggcggcgc aaccatcagt acccaatcag 780
cgaaggttga actctgacgc actgtggacg ccgagctaat aatcatcgct atggttcccg 840
gaccggcagt acttggcatt gccagcggca caaaggcaat attggcactg ggttcatctt 900
ccagctcttc cgacttgctt ttcgcctccg gtgaatcaat cgctttctgt tgcggaaaga 960
gcatccgaaa accgataaac gcgacgatta agccgcctgc aattcgcaga ccgggaatcg 1020
aaatgccaaa tgtatccatc accagttgcc cggcgtaata cgccaccatc atgatggcaa 1080
atacgtacac cgaggccatc aacgactgac gattacgttc ggcactgttc atgttgcctg 1140
ccaggccaag aaataacgcg acagttgtta atgggttagc taacggcagc aacaccacca 1200
gccccaggcc aattgcttta aacaaatcta acattggtgg ttgttatcct gtgtatctgg 1260
gttatcagcg aaaagtataa ggggtaaaca aggataaagt gtcactcttt agctagcctt 1320
gcatcgcatt gaacaaaact tgaaccgatt tagcaaaacg tggcatcggt caattcattc 1380
atttgactta tacttgcctg ggcaatatta tcccctgcaa ctaattactt gccagggcaa 1440
ctaatgtgaa aagtaccagc gatctgttca atgaaattat tccattgggt cgcttaatcc 1500
atatggttaa tcagaagaaa gatcgcctgc ttaacgagta tctgtctccg ctggatatta 1560
ccgcggcaca gtttaaggtg ctctgctcta tccgctgcgc ggcgtgtatt actccggttg 1620
aactgaaaaa ggtattgtcg gtcgacctgg gagcactgac ccgtatgctg gatcgcctgg 1680
tctgtaaagg ctgggtggaa aggttgccga acccgaatga caagcgcggc gtactggtaa 1740
aacttaccac cggcggcgcg gcaatatgtg aacaatgcca tcaattagtt ggccaggacc 1800
tgcaccaaga attaacaaaa aacctgacgg cggacgaagt ggcaacactt gagtatttgc 1860
ttaagaaagt cctgccgtaa acaaaaaaga ggtatgacga tgtccagacg caatactgac 1920
gctattacca ttcatagcat tttggactgg atcgaggaca acctggaatc gccactgtca 1980
ctggagaaag tgtcagagcg ttcgggttac tccaaatggc acctgcaacg gatgtttaaa 2040
aaagaaaccg gtcattcatt aggccaatac atccgcagcc gtaagatgac ggaaatcgcg 2100

caaaagctga aggaaagtaa cgagccgata ctctatctgg cagaacgata tggcttcgag 2160
tcgcaacaaa ctctgacccg aaccttcaaa aattactttg atgttccgcc gcataaatac 2220
cggatgacca atatgcaggg cgaatcgcgc tttttacatc cattaaatca ttacaacagc 2280
tagttgaaaa cgtgacaacg tcactgaggc aatcatgaaa ccactttcat ccgcaatagc 2340
agctgcgctt attctctttt ccgcgcaggg cgttgcggaa caaaccacgc agccagttgt 2400
tacttcttgt gccaatgtcg tggttgttcc cccatcgcag gaacacccac cgtttgattt 2460
aaatcacatg ggtactggca gtgataagtc ggatgcgctc ggcgtgccct attataatca 2520
acacgctatg tagtttgttc tggccccgac atctcggggc ttattaactt cccaccttta 2580
ccgctttacg ccaccgcaag ccaaatacat tgatatacag cccggtcata atgagcaccg 2640
cacctaaaaa ttgcagaccc gttaagcgtt catccaacaa tagtgccgca cttgccagtc 2700
ctactacggg caccagtaac gataacggtg caacccgcca ggtttcatag cgtcccagta 2760
acgtccccca gatcccataa ccaacaattg tcgccacaaa cgccagatac atcagagaca 2820
agatggtggt catatcgata gtaaccagac tgtgaatcat ggttgcggaa ccatcgagaa 2880
tcagcgaggc aacaaagaag ggaatgattg ggattaaagc gctccagatt accagcgaca 2940
tcaccgccgg acgcgttgag tgcgacatga tctttttatt gaagatgttg ccacacgccc 3000
aactaaatgc tgccgccagg gtcaacataa agccgagcat cgccacatgc tgaccgttca 3060
gactatcttc gattaacacc agtacgccaa aaatcgctaa ggcgatcccc gccaattgtt 3120
tgccatgcag tcgctccccg aaagtaaacg cgccaagcat gatagtaaaa aacgcctgtg 3180
cctgtaacac cagcgaagcc agtccagcag gcataccgaa gttaatggca caaaaaagaa 3240
aagcaaactg cgcaaaactg atggttaatc catacccccag cagcaaattc agtggtactt 3300
tcggtcgtgc gacaaaaaag atagccggaa aagcgaccag cataaagcgc aaaccggcca 3360
gcatcagcgt ggcatgttat gaagccccac tttgatgacc acaaaattta gcccccatac 3420
gaccactacc agtagcgcca acaccccatc ttttcgcgac attctaccgc ctctgaattt 3480
catcttttgt aagcaatcaa cttagctgaa tttacttttc tttaacagtt gattcgttag 3540
tcgccggtta cgacggcatt aatgcgcaaa taagtcgcta tacttcggat ttttgccatg 3600
ctatttcttt acatctctaa aacaaaacat aacgaaacgc actgccggac agacaaatga 3660
acttatccct acgacgctct accagcgccc ttcttgcctc gtcgttgtta ttaaccatcg 3720
gacgcggcgc taccgtgcca tttatgacca tttacttgag tcgccagtac agcctgagtg 3780
tcgatctaat cggttatgcg atgacaattg cgctcactat tggcgtcgtt tttagcctcg 3840
gttttggtat cctggcggat aagttcgaca agaaacgcta tatgttactg gcaattaccg 3900
ccttcgccag cggttttatt gccattactt tagtgaataa cgtgacgctg gttgtgctct 3960
tttttgccct cattaactgc gcctattctg tttttgctac cgtgctgaaa gcctggtttg 4020
ccgacaatct ttcgtccacc agcaaaacga aaatcttctc aatcaactac accatgctaa 4080
acattggctg accatcggtc cgccgctcgg cacgctgttg gta atg cag agc atc 4135
Met Gln Ser Ile
1

aat ctg ccc ttc tgg ctg gca gct atc tgt tcc gcg ttt ccc atg ctt 4183
Asn Leu Pro Phe Trp Leu Ala Ala Ile Cys Ser Ala Phe Pro Met Leu
5 10 15 20

ttc att caa att tgg gta aag cgc agc gag aaa atc atc gcc acg gaa 4231
Phe Ile Gln Ile Trp Val Lys Arg Ser Glu Lys Ile Ile Ala Thr Glu
25 30 35

aca ggc agt gtc tgg tcg ccg aaa gtt tta tta caa gat aaa gca ctg 4279
Thr Gly Ser Val Trp Ser Pro Lys Val Leu Leu Gln Asp Lys Ala Leu
40 45 50

ttg tgg ttt acc tgc tct ggt ttt ctg gct tct ttt gta agc ggc gca 4327
Leu Trp Phe Thr Cys Ser Gly Phe Leu Ala Ser Phe Val Ser Gly Ala
55 60 65

ttt gct tca tgc att tca caa tat gtg atg gtg att gct gat ggg gat 4375
Phe Ala Ser Cys Ile Ser Gln Tyr Val Met Val Ile Ala Asp Gly Asp
70 75 80

ttt gcc gaa aag gtg gtc gcg gtt gtt ctt ccg gtg aat gct gcc atg 4423
Phe Ala Glu Lys Val Val Ala Val Val Leu Pro Val Asn Ala Ala Met
85 90 95 100

```
gtg gtt acg ttg caa tat tcc gtg ggc cgc cga ctt aac ccg gct aac   4471
Val Val Thr Leu Gln Tyr Ser Val Gly Arg Arg Leu Asn Pro Ala Asn
                105                 110                 115

atc cgc gcg ctg atg aca gca ggc acc ctc tgt ttc gtc atc ggt ctg   4519
Ile Arg Ala Leu Met Thr Ala Gly Thr Leu Cys Phe Val Ile Gly Leu
            120                 125                 130

gtc ggt ttt att ttt tcc ggc aac agc ctg cta ttg tgg ggt atg tca   4567
Val Gly Phe Ile Phe Ser Gly Asn Ser Leu Leu Leu Trp Gly Met Ser
        135                 140                 145

gct gcg gta ttt act gtc ggt gaa atc att tat gcg ccg ggc gag tat   4615
Ala Ala Val Phe Thr Val Gly Glu Ile Ile Tyr Ala Pro Gly Glu Tyr
    150                 155                 160

atg ttg att gac cat att gcg ccg cca gaa atg aaa gcc agc tat ttt   4663
Met Leu Ile Asp His Ile Ala Pro Pro Glu Met Lys Ala Ser Tyr Phe
165                 170                 175                 180

tcc gcc cag tct tta ggc tgg ctt ggt gcc gcg att aac cca tta gtg   4711
Ser Ala Gln Ser Leu Gly Trp Leu Gly Ala Ala Ile Asn Pro Leu Val
                185                 190                 195

agt ggc gta gtg cta acc agc ctg ccg cct tcc tcg ctg ttt gtc atc   4759
Ser Gly Val Val Leu Thr Ser Leu Pro Pro Ser Ser Leu Phe Val Ile
            200                 205                 210

tta gcg ttg gtg atc att gct gcg tgg gtg ctg atg tta aaa ggg att   4807
Leu Ala Leu Val Ile Ile Ala Ala Trp Val Leu Met Leu Lys Gly Ile
        215                 220                 225

cga gca aga ccg tgg ggg cag ccc gcg ctt tgt tga tttaagtcga        4853
Arg Ala Arg Pro Trp Gly Gln Pro Ala Leu Cys  *
    230                 235

acacaataaa gatttaattc agccttcgtt taggttacct ctgctaatat ctttctcatt 4913
gagatgaaaa ttaaggtaag cgaggaaaca caccacacca taaacggagg caaataatgc 4973
tgggtaatat gaatgttttt atggccgtac tgggaataat tttatttttct ggttttctgg 5033
ccgcgtattt cagccacaaa tgggatgact aatgaacgga gataatccct cacctaaccg 5093
gccccttgtt acagttgtgt acaaggggcc tgatttttat gacggcgaaa aaaaaccgcc 5153
agtaaaccgg cggtgaatgc ttgcatggat agatttgtgt tttgctttta cgctaacagg 5213
cattttcctg cactgataac gaatcgttga cacagtagca tcagttttct caatgaatgt 5273
taaacggagc ttaaactcgg ttaatcacat tttgttcgtc aataaacatg cagcgatttc 5333
ttccggtttg cttaccctca tacattgccc ggtccgctct tccaatgacc acatccagag 5393
gctcttcagg aaatgcgcga ctcacacctg ctgtcacggt aatgttgata tgcccttcag 5453
aatgtgtgat ggcatggtta tcgactaact ggcaaattct gacacctgca cgacatgctt 5513
cttcatcatt agccgctttg acaataatga taaattcttc gcccccgtag cgataaaccg 5573
tttcgtaatc acgcgtccaa ctggctaagt aagttgccag ggtgcgtaat actacatcgc 5633
cgattaaatg cccgtagtat cattaaccaa tttaaatcgg tcaatatcca acaacattaa 5693
ataaagattc agaggctcag cgttgcgtaa ctgatgatca aaggattcat caagaacccg 5753
acgacccggc aatcccgtca aaacatccat attgctacgg atcgtcagca aataaatttt 5813
gtaatcggtt aatgccgcag taaaagaaag caacccctcc tgaaaggcgt cgaaatgcgc 5873
gtcctgccag tgattttcaa caatagccag cattaattcc cgaccacagt tatgcatatg 5933
ttgatgggca gaatccatta gccgaacgta aggtaattca tcgttatcga gtggccccag 5993
atgatcaatc caccgaccaa actggcacag tccataagaa tggttatccg ttatttctgg 6053
cttactggca tctctcgcga ccacgctgtg aaacatactc accagccact ggtagtgggc 6113
atcgatagcc ttattgagat ttaacaagat ggcatcaatt tccgttgtct tcttgatcat 6173
tgccactcct ttttcacagt tccttgtgcg cgctattcta acgagagaaa agcaaaatta 6233
cgtcaatatt ttcatagaaa tccgaagtta tgagtcatct ctgagataac attgtgattt 6293
aaaacaaaat cagcggataa aaaagtgttt aattctgtaa attacctctg cattatcgta 6353
```

```
aataaaagga tgacaaatag cataacccaa taccctaatg gcccagtagt tcaggccatc 6413
aggctaattt atttttattt ctgcaaatga gtgacccgaa cgacggccgg cgcgcttttc 6473
ttatccagac tgccactaat gttgatcatc tggtccggct gaacttctcg tccatcaaag 6533
acggccgcag gaataacgac attaatttca ccgctcttat cgcgaaaaac gtaacggtcc 6593
tctcctttgt gagaaatcaa attaccgcgt agtgaaaccg aagcgccatc gtgcatggtt 6653
tttgcgaaat caacggtcat ttttttttgca tcatcggttc cgcgatagcc atcttctatt 6713
gcatgaggcg gcggtggcgc tgcatcctgt tttaaaccgc cctggtcatc tgccaacgca 6773
taaggcatga caagaaaact tgctaataca atggcctgaa atttcatact aactccttaa 6833
ttgcgtttgg tttgacttat taagtctggt tgctattttt ataattgcca aataagaata 6893
ttgccaattg ttataaggca tttaaaatca gccaactagc tgtcaaatat acagagaatt 6953
taactcacta aagttaagaa gattgaaaag tcttaaacat attttcagaa taatcggatt 7013
tatatgtttg aaaattatta tattggacga gcatacagaa aaagcaaatc acctttacat 7073
ataaaagcgt ggacaaaaaa cagtgaacat taatagagat aaaattgtac aacttgtaga 7133
taccgatact attgaaaacc tgacatccgc gttgagtcaa agacttatcg cggatcaatt 7193
acgcttaact accgccgaat catgcaccgg cggtaagttg gctagcgccc tgtgtgcagc 7253
tgaagataca cccaaatttt acggtgcagg ctttgttact ttcaccgatc aggcaaagat 7313
gaaaatcctc agcgtaagcc agcaatctct tgaacgatat tctgcggtga gtgagaaagt 7373
ggcagcagaa atggcaaccg gtgccataga gcgtgcggat gctgatgtca gtattgccat 7433
taccggctac ggcggaccgg agggcggtga agatggtacg ccagcgggta ccgtctggtt 7493
tgcgtggcat attaaaggcc agaactacac tgcggttatg cattttgctg gcgactgcga 7553
aacggtatta gctttagcgg tgaggtttgc cctcgcccag ctgctgcaat tactgctata 7613
accaggctgg cctggcgata tctcaggcca gccattggtg gtgtttatat gttcaagcca 7673
cgatgttgca gcatcggcat aatcttaggt gccttaccgc gccattgtcg atacaggcgt 7733
tccagatctt cgctgttacc tctggaaagg atcgcctcgc gaaaacgcag cccattttca 7793
cgcgttaatc cgccctgctc aacaaaccac tgataaccat catcggccaa catttgcgtc 7853
cacagataag cgtaataacc tgcag                                      7878
```

<210> 2
<211> 239
<212> PRT
<213> Echerichia coli

<400> 2

```
Met Gln Ser Ile Asn Leu Pro Phe Trp Leu Ala Ala Ile Cys Ser Ala
 1               5                  10                  15
Phe Pro Met Leu Phe Ile Gln Ile Trp Val Lys Arg Ser Glu Lys Ile
            20                  25                  30
Ile Ala Thr Glu Thr Gly Ser Val Trp Ser Pro Lys Val Leu Leu Gln
        35                  40                  45
Asp Lys Ala Leu Leu Trp Phe Thr Cys Ser Gly Phe Leu Ala Ser Phe
    50                  55                  60
Val Ser Gly Ala Phe Ala Ser Cys Ile Ser Gln Tyr Val Met Val Ile
65                  70                  75                  80
Ala Asp Gly Asp Phe Ala Glu Lys Val Val Ala Val Val Leu Pro Val
                85                  90                  95
Asn Ala Ala Met Val Val Thr Leu Gln Tyr Ser Val Gly Arg Arg Leu
            100                 105                 110
Asn Pro Ala Asn Ile Arg Ala Leu Met Thr Ala Gly Thr Leu Cys Phe
        115                 120                 125
Val Ile Gly Leu Val Gly Phe Ile Phe Ser Gly Asn Ser Leu Leu Leu
    130                 135                 140
Trp Gly Met Ser Ala Ala Val Phe Thr Val Gly Glu Ile Ile Tyr Ala
145                 150                 155                 160
Pro Gly Glu Tyr Met Leu Ile Asp His Ile Ala Pro Pro Glu Met Lys
                165                 170                 175
Ala Ser Tyr Phe Ser Ala Gln Ser Leu Gly Trp Leu Gly Ala Ala Ile
            180                 185                 190
Asn Pro Leu Val Ser Gly Val Val Leu Thr Ser Leu Pro Pro Ser Ser
        195                 200                 205
Leu Phe Val Ile Leu Ala Leu Val Ile Ile Ala Ala Trp Val Leu Met
    210                 215                 220
```

```
Leu Lys Gly Ile Arg Ala Arg Pro Trp Gly Gln Pro Ala Leu Cys
225                 230                 235


<210> 3
<211> 870
<212> DNA
<213> Echerichia coli

<220>
<221> CDS
<222> (1)...(870)

<400> 3
atg gat cag gcc ggc att att cgc gac ctt tta atc tgg ctg gaa ggt   48
Met Asp Gln Ala Gly Ile Ile Arg Asp Leu Leu Ile Trp Leu Glu Gly
 1               5                  10                  15

cat ctg gat cag ccc ctg tcg ctc gac aat gta gcg gcg aaa gca ggt   96
His Leu Asp Gln Pro Leu Ser Leu Asp Asn Val Ala Ala Lys Ala Gly
            20                  25                  30

tat tcc aag tgg cac tta cag aga atg ttt aaa gat gtc act ggc cat   144
Tyr Ser Lys Trp His Leu Gln Arg Met Phe Lys Asp Val Thr Gly His
        35                  40                  45

gct att ggc gcg tat att cgt gct cgt cgt ttg tcg aaa tcg gcg gtc   192
Ala Ile Gly Ala Tyr Ile Arg Ala Arg Arg Leu Ser Lys Ser Ala Val
    50                  55                  60

gca cta cgc ctg act gcg cgt ccg att ctg gac atc gcg ctg caa tac   240
Ala Leu Arg Leu Thr Ala Arg Pro Ile Leu Asp Ile Ala Leu Gln Tyr
65                  70                  75                  80

cgc ttc gac tct caa cag aca ttt acc cgc gca ttc aag aag cag ttt   288
Arg Phe Asp Ser Gln Gln Thr Phe Thr Arg Ala Phe Lys Lys Gln Phe
                85                  90                  95

gcc cag act cct gca ctt tac cgc cgt tct cct gaa tgg agc gcc ttt   336
Ala Gln Thr Pro Ala Leu Tyr Arg Arg Ser Pro Glu Trp Ser Ala Phe
            100                 105                 110

ggt att cgc ccg ccg cta cgc ctg ggt gaa ttc act atg cca gag cac   384
Gly Ile Arg Pro Pro Leu Arg Leu Gly Glu Phe Thr Met Pro Glu His
        115                 120                 125

aaa ttt gtc acc ctg gaa gat acg ccg ctg att ggt gtt acc cag agc   432
Lys Phe Val Thr Leu Glu Asp Thr Pro Leu Ile Gly Val Thr Gln Ser
    130                 135                 140

tac tcc tgt tcg ctg gag caa atc tct gat ttc cgc cat gaa atg cgt   480
Tyr Ser Cys Ser Leu Glu Gln Ile Ser Asp Phe Arg His Glu Met Arg
145                 150                 155                 160

tat cag ttc tgg cac gat ttt ctc ggc aac gcg ccg acc att ccg ccg   528
Tyr Gln Phe Trp His Asp Phe Leu Gly Asn Ala Pro Thr Ile Pro Pro
                165                 170                 175

gtg ctc tac ggc ctg aat gaa acg cgt ccg agt cag gat aaa gac gac   576
Val Leu Tyr Gly Leu Asn Glu Thr Arg Pro Ser Gln Asp Lys Asp Asp
            180                 185                 190
```

```
gag caa gag gta ttc tat acc acc gcg tta gcc cag gat cag gca gat   624
Glu Gln Glu Val Phe Tyr Thr Thr Ala Leu Ala Gln Asp Gln Ala Asp
        195                 200                 205

ggc tat gta ctg acg ggg cat ccg gtg atg ctg cag ggc ggc gaa tat   672
Gly Tyr Val Leu Thr Gly His Pro Val Met Leu Gln Gly Gly Glu Tyr
    210                 215                 220

gtg atg ttt acc tat gaa ggt ctg gga acc ggc gtg cag gag ttt atc   720
Val Met Phe Thr Tyr Glu Gly Leu Gly Thr Gly Val Gln Glu Phe Ile
225                 230                 235                 240

ctg acg gta tac gga acg tgc atg cca atg ctc aac ctg acg cgc cgt   768
Leu Thr Val Tyr Gly Thr Cys Met Pro Met Leu Asn Leu Thr Arg Arg
                245                 250                 255

aaa ggt cag gat att gag cga tac tac ccg gca gaa gat gcc aaa gcg   816
Lys Gly Gln Asp Ile Glu Arg Tyr Tyr Pro Ala Glu Asp Ala Lys Ala
            260                 265                 270

gga gat cgc cca att aat cta cgc tgt gaa ctg ctg att ccg atc cgt   864
Gly Asp Arg Pro Ile Asn Leu Arg Cys Glu Leu Leu Ile Pro Ile Arg
        275                 280                 285

cgt taa                                                          870
Arg  *
```

<210> 4
<211> 289
<212> PRT
<213> Echerichia coli

<400> 4

```
Met Asp Gln Ala Gly Ile Ile Arg Asp Leu Leu Ile Trp Leu Glu Gly
 1               5                  10                  15
His Leu Asp Gln Pro Leu Ser Leu Asp Asn Val Ala Ala Lys Ala Gly
            20                  25                  30
Tyr Ser Lys Trp His Leu Gln Arg Met Phe Lys Asp Val Thr Gly His
        35                  40                  45
Ala Ile Gly Ala Tyr Ile Arg Ala Arg Arg Leu Ser Lys Ser Ala Val
    50                  55                  60
Ala Leu Arg Leu Thr Ala Arg Pro Ile Leu Asp Ile Ala Leu Gln Tyr
65                  70                  75                  80
Arg Phe Asp Ser Gln Gln Thr Phe Thr Arg Ala Phe Lys Lys Gln Phe
                85                  90                  95
Ala Gln Thr Pro Ala Leu Tyr Arg Arg Ser Pro Glu Trp Ser Ala Phe
            100                 105                 110
Gly Ile Arg Pro Pro Leu Arg Leu Gly Glu Phe Thr Met Pro Glu His
        115                 120                 125
Lys Phe Val Thr Leu Glu Asp Thr Pro Leu Ile Gly Val Thr Gln Ser
    130                 135                 140
Tyr Ser Cys Ser Leu Glu Gln Ile Ser Asp Phe Arg His Glu Met Arg
145                 150                 155                 160
Tyr Gln Phe Trp His Asp Phe Leu Gly Asn Ala Pro Thr Ile Pro Pro
                165                 170                 175
Val Leu Tyr Gly Leu Asn Glu Thr Arg Pro Ser Gln Asp Lys Asp Asp
            180                 185                 190
Glu Gln Glu Val Phe Tyr Thr Thr Ala Leu Ala Gln Asp Gln Ala Asp
        195                 200                 205
```

```
Gly Tyr Val Leu Thr Gly His Pro Val Met Leu Gln Gly Gly Glu Tyr
    210                 215                 220
Val Met Phe Thr Tyr Glu Gly Leu Gly Thr Gly Val Gln Glu Phe Ile
225                 230                 235                 240
Leu Thr Val Tyr Gly Thr Cys Met Pro Met Leu Asn Leu Thr Arg Arg
                245                 250                 255
Lys Gly Gln Asp Ile Glu Arg Tyr Tyr Pro Ala Glu Asp Ala Lys Ala
            260                 265                 270
Gly Asp Arg Pro Ile Asn Leu Arg Cys Glu Leu Leu Ile Pro Ile Arg
        275                 280                 285
Arg
```